

FACILITIES INFORMATION MANAGEMENT SYSTEM

PHASE I. METRICS VALIDATION STANDARD

PHASE II. METRICS VALIDATION PROCESS

September 2004

Odoi Associates, Inc.
FACILITY & CONSTRUCTION MANAGEMENT SERVICES

LMI
GOVERNMENT CONSULTING

FACILITIES INFORMATION MANAGEMENT SYSTEM

PHASE I. METRICS VALIDATION STANDARD

PHASE II. METRICS VALIDATION PROCESS

REPORT OD401C1

Odoi Associates Inc.

Robert E. Ward

Brent L. Vanderpool

LMI

Richard A. Skulte

Jonathan P. Adams

Robert L. Crosslin

Anna M. Wallace

Gerald W. Westerbeck

September 2004

Odoi Associates, Inc.
FACILITY & CONSTRUCTION MANAGEMENT SERVICES

LMI
GOVERNMENT CONSULTING

LEGAL NOTICE :

THE VIEWS, OPINIONS, AND FINDINGS CONTAINED IN THIS REPORT ARE THOSE OF LMI AND SHOULD NOT BE CONSTRUED AS AN OFFICIAL AGENCY POSITION, POLICY, OR DECISION, UNLESS SO DESIGNATED BY OTHER OFFICIAL DOCUMENTATION.

©LMI 2004

DOE Facilities Information Management System

OD401C1/SEPTEMBER 2004

Executive Summary

The Department of Energy, the steward of the nation's energy resources, has real property holdings of over 2.4 million acres of land and 120 million square feet of buildings. Its Facilities Information Management System (FIMS) database provides ready access to DOE data and contains information with which to generate an annual report summarizing the size and cost of real property holdings. Complete and accurate FIMS information is critical for asset planning, programming, budgeting, and management.

Due to the large volume of data within FIMS and the decentralized nature of data entry, data accuracy is unknown. DOE's Office of Engineering and Construction Management (OECM) asked Odoi Associates, Inc., and LMI to develop a standard, a process, and an oversight method for validating the accuracy of its FIMS data. This report provides the results of that effort: a standard with which to validate the accuracy of three key FIMS metrics and the results of a pilot test of the process. The three key FIMS metrics are

- ◆ replacement plant value (RPV),
- ◆ asset condition index (ACI), and
- ◆ asset utilization index (AUI).

Since it would be too costly and time consuming for DOE to validate the FIMS data for each asset, we developed a statistical standard for determining the accuracy of each metric at the site level, program level, and DOE-wide. Using an allowable error of ± 10 percent, we recommend the following:

- ◆ *Validate all three metrics using the same sample.* Since the components of each metric are related, and often exist in multiple indexes, the same sample can be used to validate all three metrics. The three metrics interrelate on the basis of common data elements within their respective formulas. Gross square footage is a key data element in calculating RPV and AUI, and potentially a large source of error. Since RPV is used to calculate ACI, gross square footage is also a key element of that calculation.

-
- ◆ *Use a 90 percent level of confidence for statistical sampling and analysis.* A minimum site sample size of up to 25 is required to achieve at least 90 percent level of confidence.¹ A minimum site sample size of up to 36 would be required to achieve a 95 percent level of confidence—a 44 percent increase in sampling size and likely a commensurate level of required effort. For the purposes of budgetary decisions, the achievable benefit does not warrant the additional level of effort.
 - ◆ *Use simple random sampling at first.* About 85 percent of all facilities are less than 10,000 square feet, and about 50 percent have replacement costs less than \$100,000. Therefore, simple random sampling is sufficient and should be used. Future validations can weight sampling by facility square footage and then compare the results with those of simple random sampling.
 - ◆ *Investigate all outliers.* The 20 to 30 outliers, defined as high gross square footage with low RPV or low gross square footage with high RPV, should be added to the random sample in order to understand whether data for these facilities are accurate.
 - ◆ *Focus the validation on the incidence or frequency of errors in each metric and the main sources or causes of those errors.* DOE does not use a mission criticality index to segregate or prioritize the facilities and their metrics. Focusing on the frequency of errors first, providing the most objective learning process in understanding the main sources of error, best serves the validation.
 - ◆ *Sample at the site level.* Several sites contain facilities belonging to more than one program. However, facilities at a site are typically operated and maintained according to a single set of policies and procedures. Sampling at the site level (up to 25 facilities at each site) allows valid statistical inferences to be made regarding the frequency of errors at each site. And, given the total number of facilities sampled for each program across all sites, this sampling method also allows valid statistical inferences regarding the aggregate frequency of errors for a program as well as for all of DOE.

We tested the method at two pilot sites—National Energy Technology Laboratory (Fossil Energy program) in Morgantown, WV, and Brookhaven National Laboratory (Science program) in Brookhaven, NY. On the basis of the pilot, we add the following recommendations:

- ◆ *Use a “desk validation” approach.* A desk validation should be used, supplemented by a small “sample of the sample” for physical validation, during the first round of full-scale implementation. The desk validation uses available source documents (AutoCad for gross square footage, etc.),

¹ See Appendix C-1, table C-1 for list of required sample size by site population.

and the pilot sample of the sample verified that it accurately reflects FIMS. Continuing the sample of the sample during the first full-scale FIMS validation provides non-site members of the evaluation team with a first-hand understanding of the way the site applies facility-related policies.

- ◆ *Incorporate the validation process into the FIMS web application.* The validation standard and process should be implemented as a module of the web-based version of FIMS. The random sampling of facilities, pre-populating the data collection form, and aggregating the results by site, program, and DOE can be automated as part of this database application.
- ◆ *Have an independent party lead the validation team.* A site-independent member—from headquarters, the site’s servicing center, or another site from a different program—should lead the validation team at each site. Headquarters should develop a short FIMS validation training program for these team leads.
- ◆ *Develop policy guidance.* Examine and document best practices certain sites or programs already perform. These best practices can then be shared or formalized into policy guidance that defines the intended meaning, interpretation, and application of the key metrics within FIMS.
- ◆ *Identify methods for OECM oversight.* Develop built-in controls, such as logic traps and exception reports, to improve quality.

The FIMS validation standard and process will allow DOE to improve the accuracy of its data and support key real property decisions.

Contents

Preface.....	ix
--------------	----

Phase I. Metrics Validation Standard

CHAPTER 1 INTRODUCTION	1-1-1
Background	1-1-1
Objective	1-1-1
Structure	1-1-2
CHAPTER 2 VALIDATION STANDARD	1-2-1
Information Sources	1-2-1
Statistical Approach	1-2-1
Calculations	1-2-3
Square Footage Weighting	1-2-4
Facility Selection	1-2-8
CHAPTER 3 FIMS METRICS	1-3-1
Replacement Plant Value	1-3-1
Asset Condition Index	1-3-2
Asset Utilization Index	1-3-3
CHAPTER 4 CONCLUSIONS AND RECOMMENDATIONS	1-4-1
Validate all three metrics using the same sample	1-4-1
Use a 90 percent level of confidence for statistical sampling and analysis....	1-4-1
Use simple random sampling at first	1-4-1
Sample all outliers.	1-4-1
Focus the validation on the incidence or frequency of errors in each metric and the main sources or causes of those errors.	1-4-2
Sample each program separately at each site	1-4-2

Phase I. Metrics Validation Process

CHAPTER 1 INTRODUCTION	2-1-1
Objective	2-1-1
Structure	2-1-1

CHAPTER 2 PILOT VALIDATION	2-2-1
Purpose	2-2-1
Expectations	2-2-1
CHAPTER 3 VALIDATION PROCESS.....	2-3-1
Overview	2-3-1
Initialization.....	2-3-1
Validation.....	2-3-4
Evaluation and Reporting	2-3-14
CHAPTER 4 PILOT OBSERVATIONS.....	2-4-1
Overview	2-4-1
National Energy Technology Laboratory	2-4-1
Brookhaven National Laboratory	2-4-6
CHAPTER 5 CONCLUSIONS AND RECOMMENDATIONS	2-5-1
Overview	2-5-1
Online Sampling, Data Collection, and Results	2-5-1

Appendix A. References

Appendix B. Best Practices

Appendix C. Calculated Sample Sizes

Appendix D. FIMS Validation Handbook

Figures

Figure 1-2-1. Frequency Distribution of Gross Square Footage for All Facilities	1-2-5
Figure 1-2-2. Frequency Distribution of Gross Square Footage for EM Program.....	1-2-5
Figure 1-2-3. Frequency Distribution of Gross Square Footage for NNSA Program.....	1-2-6
Figure 1-2-4. Frequency Distribution of RPV for All Facilities	1-2-7
Figure 1-2-5. RPV vs. Gross Square Footage for All Facilities	1-2-7

Tables

Table 2-3-1. RPV Data Elements.....	2-3-5
-------------------------------------	-------

Preface

The Department of Energy's Office of Engineering and Construction Management asked Odoi Associates, Inc., and LMI to develop a validation standard and process for its Facilities Information Management System (FIMS) data. To do so, we first developed a statistical standard for validating the accuracy of key FIMS metrics, and then we created a process for the validation. Finally, we pilot tested the standard and process.

DOE accepted the standard and process and selected two sites for the pilot test: the National Energy Technology Laboratory (Morgantown, WV) and Brookhaven National Laboratory (Brookhaven, NY). LMI assisted DOE in conducting the pilot test, evaluated the results, and incorporated feedback and lessons learned into the final FIMS validation standard and process report.

This report documents the two efforts. The first part describes Phase I, developing the statistical standard, and the second describes Phase II, the validation process and pilot test.

FACILITIES INFORMATION MANAGEMENT SYSTEM

PHASE I. METRICS VALIDATION STANDARD

Odoi Associates Inc.

Robert E. Ward

Brent L. Vanderpool

LMI

Richard A. Skulte

Jonathan P. Adams

Robert L. Crosslin

Anna M. Wallace

Gerald W. Westerbeck

SEPTEMBER 2004

Chapter 1

Introduction

The Department of Energy's Office of Engineering and Construction Management (OECM) asked LMI—through Odoi Associates, Inc. (OAI)—to develop a statistical standard for validating the accuracy of its Facilities Information Management System (FIMS) data. This part of our report, Phase I, describes the standard we developed.

BACKGROUND

DOE Order 430.1B requires the department to establish a holistic, performance-based approach to real property life-cycle asset management and implement it by September 30, 2004.¹ This new directive integrates various management elements that will improve the management of DOE's real property assets. One specific element of this broader objective is the requirement to develop a process to validate the completeness and quality of FIMS data, which links asset planning, programming, budgeting, and evaluation to the department's multifaceted missions. OECM has made the decision to develop a validation process to fulfill this requirement and comply with *The President's Management Agenda*.

FIMS assists DOE in managing its corporate physical assets. As the corporate real property database, it manages all real property, including land and anything permanently affixed to it. DOE, the steward of the nation's energy resources, has real property holdings of over 2.4 million acres of land and 120 million square feet of buildings. FIMS provides ready access to DOE data and contains information with which to generate an annual report summarizing the size and cost of real property holdings. Complete and accurate information is critical for managing facilities and reporting to the General Services Administration (GSA), Office of Management and Budget (OMB), Congress, and the taxpayers.

OBJECTIVE

The objective of Phase I was to develop a statistical standard for validating the accuracy of three key FIMS metrics at the site and program levels:²

- ◆ Replacement plant value (RPV)

¹ U.S. Department of Energy, DOE Order 430.1B, *Real Property Asset Management (RPAM)*, (Washington, DC: September 24, 2003).

² A site is a group of DOE facilities in one geographic location. DOE programs include the National Nuclear Security Administration (NNSA), Science, Nuclear Energy, Fossil Energy, Energy Efficiency, and Renewable Energy.

-
- ◆ Asset condition index (ACI)
 - ◆ Asset utilization index (AUI).

As part of this objective, we recommend a definition of accuracy, level of statistical confidence, and asset sampling guidelines. We also identify the primary causes or sources of error for these metrics when inaccuracies occur.

STRUCTURE

The remainder of this part of our report comprises the following chapters:

- ◆ In Chapter 2, we describe development of sampling guidelines, documentation of assumptions, and accuracy definitions.
- ◆ Chapter 3, we give an overview of the FIMS database and detail the three metrics and their sub-elements, including potential sources and causes of inaccuracies.
- ◆ In Chapter 4, we provide our conclusions and recommendations.
- ◆ In the references section, we list some of the documents used in the preparation of this report.
- ◆ In the appendixes, we provide supporting information, including statistical calculations with sampling size results by site and program.

Chapter 2

Validation Standard

INFORMATION SOURCES

In developing a validation standard for the FIMS database, we considered the following sources of information:

- ◆ *FIMS database.* With assistance from OECM and OAI, we reviewed the scope and hierarchy of the database, values and ranges of key data elements, standard report outputs, and custom data queries.
- ◆ *FIMS website and official documentation.* We reviewed the FIMS users guide, training manual, data dictionary, and other related documentation provided by OECM.
- ◆ *OECM personnel.* We met with OECM to review in detail each metric and related data elements and to understand the overall uses of the FIMS database and underlying implications of accuracy.
- ◆ *Best practices.* We reviewed the data validation practices of several public and private-sector organizations to align our approach, where appropriate, with the best practices of those organizations. Organizations we looked at include NAVFAC, ETA, USACE, EPA, and RTI.¹

STATISTICAL APPROACH

To develop a statistical standard for validating the metrics, we considered the following:

- ◆ Their uses
- ◆ The level of error allowed or tolerated in the measurement of the metrics
- ◆ The desired level of confidence in the estimated accuracy of the metrics, which depends (in part) on their uses.

DOE primarily uses the metrics in budget formulation and budget execution related to facility repair and replacement. Although important, these uses are not urgent compared with, say, determining whether exposure to a given substance will cause a serious illness or disease. In the latter case, a level of allowable (tol-

¹ NAVFAC = Naval Facilities Engineering Command; ETA = Employment and Training Administration, USACE = U.S. Army Corps of Engineers, EPA = U.S. Environmental Protection Agency; and RTI = Research Triangle Institute.

erable) error might be 1 percent or less, and the desired level of confidence in the accuracy of the measurement might be 95 percent, or even 99 percent.

In addition, for the purposes of the validation process, it is only necessary to determine whether a site's and a program's reported metrics are "accurate or inaccurate within a level of tolerance," and not the "total estimated amount of inaccuracy" in the metrics. Given the budgetary use of the site- and program-level metrics, we recommended

- ◆ an allowable (tolerable) error of 10 percent, and
- ◆ a level of confidence of 90 percent that the estimated accuracy of the metric will be within the tolerable level.

As a result, we developed a "simple random sampling" plan that allows DOE to make the following type of statements at the end of the validation process:

- ◆ *Site level.* For a given program at a given site, we are 90 percent confident that x percent of asset RPVs are accurate within ± 10 percent. (A similar statement can be made for the ACI and AUI metrics for each program at each site.)
- ◆ *Program level.* For a given program overall, we are 90 percent confident that x percent of asset RPVs are accurate within ± 10 percent. (A similar statement can be made for the ACI and AUI metrics for each program overall.)

These statements are with respect to "facilities," not "square footage." In other words, under the simple random sampling plan, each facility in each program at a given site has an equally likely chance of selection for validation, regardless of size. It may be desirable to weight the sample selection process by a measure of square footage, giving larger facilities a higher chance of being included in the sample and ensuring representation across the size distribution of facilities.² In either case, the sample size for each separate program at each site varies between 1 and 25, depending on the total number of facilities in each program at the site. We use the same sample of each program at each site to determine the accuracy of each of the three metrics and to calculate the estimated accuracy of the metrics by program and site.

In Phase II, we develop a process for determining "root causes" of errors identified during the validation. This process focuses on determining, for each facility found to be "outside the allowable error of 10 percent" for a given metric, the probable sources of the error. We develop guidelines for using this information to conduct "root cause analysis" to identify the major sources of error and to aid in developing error-mitigation strategies.

² We recommend a method in the second part of this report.

CALCULATIONS

The approach that we recommend—determining whether the metric for a facility is accurate or inaccurate—results in a situation characterized in statistics as a “binomial distribution.” In other words, a portion, p , of the population has some attribute, and the remainder of the population, $1-p$, or q , does not.³ In this case, the attribute is “accuracy of the metric for a particular facility.”

We have no strong reason to suspect that the accuracy of any one metric significantly differs from the accuracy of another. In fact, each one contains common items (for example, square feet), so the accuracy of each metric for a given facility should be similar and related to the accuracy of the other metrics for that facility. For these reasons, our calculation of sample size for each program and site uses a single estimate of p (and therefore of q) equal to 0.9. In other words, we expect that the accuracy of each metric, for each program and at each site, is approximately 90 percent.

As discussed previously, the amount of “allowable error,” L , in a particular facility’s metric is 10 percent, therefore the level of confidence desired in the accuracy of a metric for a program at a site is 90 percent (“0.90 confidence probability”). Using the formulas for sampling from a binomial distribution,⁴ we find that these assumptions yield the following sample size calculations for each separate program at each site:

$$n = \frac{2.7 pq}{L^2}, \quad [1]$$

or, since we are dealing with relatively small populations per site, we use a finite population correction factor:

$$n' = \frac{n}{\left(\frac{1+n}{N} \right)}, \quad [2]$$

where N is the total population of facilities of that program type at the installation, and n exceeds 10 percent of N .

Application of these formulas results in sample sizes (n or n') that vary from 1 to 25 for a program at a given site, depending on the total (N) of facilities of that particular program type at that site. Table A-1 shows the required sample size for each program at a site, using the above formulas, for all values of N up to 243⁵ as well as expected sample sizes by program by site (see Appendix C). If a 95

³ George W. Snedecor and William G. Cochran, *Statistical Methods*, sixth edition (Ames, Iowa: The Iowa State University Press, 1967), p. 202.

⁴ See Note 3, pp. 516–518.

⁵ After which the same size remains constant at 25.

percent level of confidence is desired, the maximum sample size for each program at each site increases to 36, which would result in validating up to 44 percent more facilities in each program at each site than required by a 90 percent level of confidence. The 5 percent increase in level of confidence does not justify the increased time and effort in this situation on the basis of our initial finding regarding the quality of data for facilities considered non-life-threatening.

We can “roll up” the results to the DOE level for each program for each metric. We simply put all sampled facilities in a given program across all sites into a single group and use these observations to calculate the accuracy (frequency of error and non-error) of the metrics for that program at the DOE level.⁶

SQUARE FOOTAGE WEIGHTING

As discussed previously, the above statements and approach are with respect to “facilities,” not “square footage.” In other words, with simple random sampling, each facility in each program at a given site has an equally likely chance of being selected for validation. It may be desirable to weight the sample selection process by a measure of square footage, giving larger facilities a higher chance of being included in the sample at a future date.

We analyzed the size (square footage) distribution of facilities for the individual programs and for all programs at the two proposed pilot sites; the size is not normally distributed. Our analysis indicates that, for most programs, the vast majority (about 85 percent) of facilities are relatively small (less than 10,000 square feet). Figures 1-2-1 through 1-2-3 show frequency histograms of gross square footage for all programs and separately for the NNSA and EM programs (which account for two-thirds of DOE facilities in FIMS).⁷ Eight-five percent of the facilities are small and were not included in the weighting process.

Simple random sampling (each facility has an equal chance of being selected for validation) results in very few large facilities. This approach and outcome is desirable if the focus is on the “incidence” (frequency) of error in FIMS. It is also desirable if mission criticality of facilities is unrelated, or negatively correlated, with size of facilities. For these reasons, we recommend that this first effort to validate FIMS use the simple random sampling approach, which provides an accurate estimate of the incidence or frequency of error and causes of error in the relatively new FIMS program. It also allows for policies and actions to reduce the frequency of error across all facilities.

⁶ We are sampling to determine the accuracy of each program at each site and therefore have up to 25 sampled facilities in each program that has facilities at a given site. With respect to each program overall in DOE, this represents an “over-sampling” to achieve the required confidence as to the accuracy of each program in total.

⁷ Due to the wide range of values for gross square footage, we chose to display up to 285,000 square feet on the x-axis. The actual high range is about 2.8 million; if we displayed that scale, the graphs would be unreadable.

Figure 1-2-1. Frequency Distribution of Gross Square Footage for All Facilities

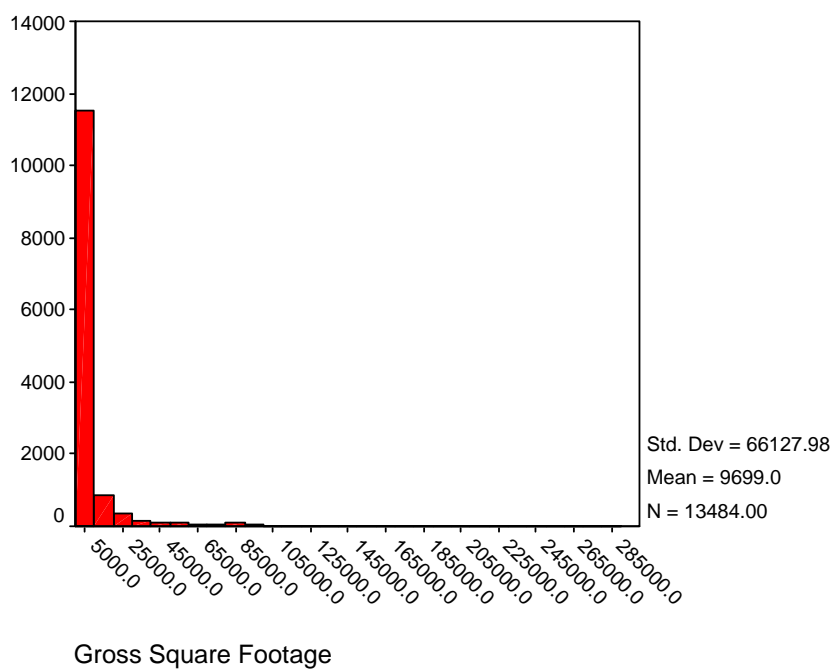


Figure 1-2-2. Frequency Distribution of Gross Square Footage for EM Program

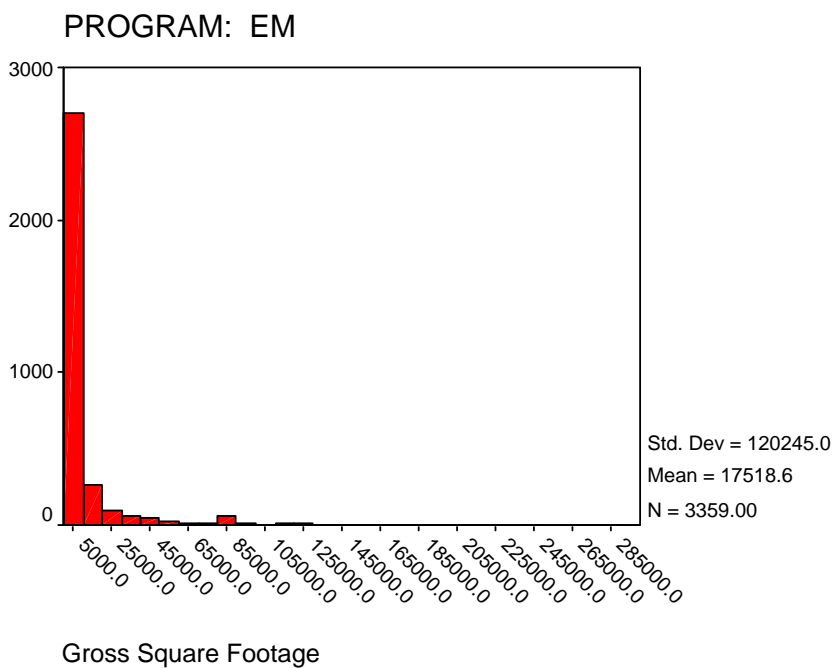
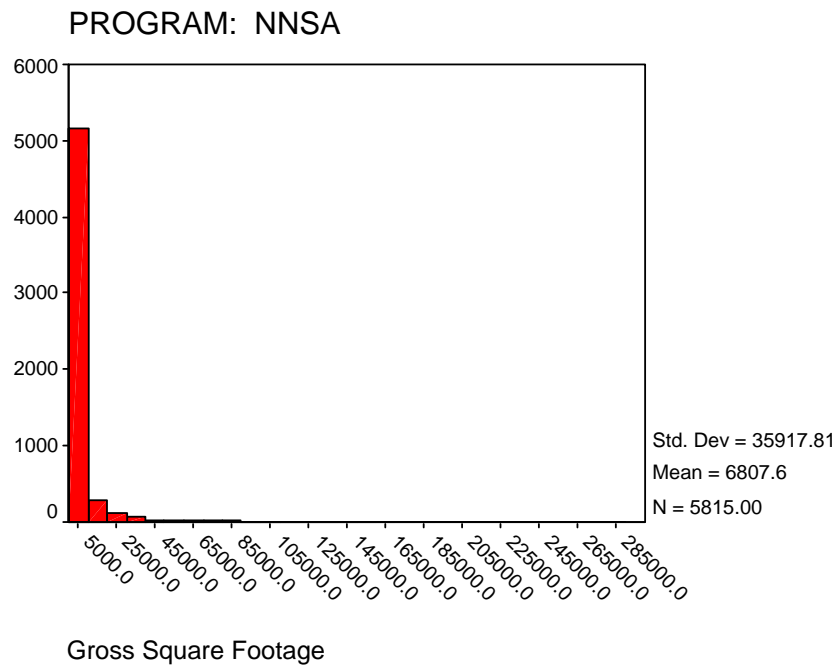


Figure 1-2-3. Frequency Distribution of Gross Square Footage for NNSA Program



Since square footage is an important driver of RPV, it is not surprising that about 80 percent of all replacement values are less than \$100,000, as can be seen from Figure 1-2-4 (a histogram of the frequency distribution of RPVs).⁸ The scatter-gram of RPV against gross square footage in Figure 1-2-5 shows that the overwhelming majority of DOE facilities are concentrated together in a range that can be characterized as “small with low replacement cost.” Together, these figures indicate that weighting the sample by square footage would not give different results than simple random sampling. The results would be the same regardless of their use. However, in future validation efforts (after correcting policies and procedures), we recommend that DOE use a weighted random sampling technique on one occasion, weighting each facility by its reported gross square footage as a percentage of total reported square footage for that program at that site. We do not expect a significant difference in results although the technique may provide a different estimate, compared with simple random sampling, of the accuracy of the metrics when used for budgetary purposes. This technique will allow the error rates of larger facilities to be compared with those of the overall population.

⁸ Again, for readability, we chose to display only a portion of the x-axis scale. The high end RPV range is about \$1.6 billion.

Figure 1-2-4. Frequency Distribution of RPV for All Facilities

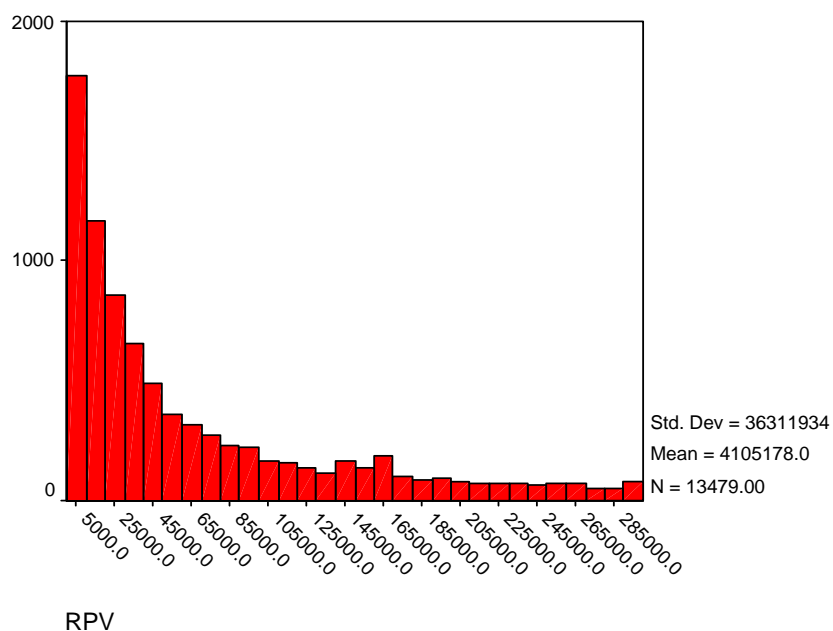
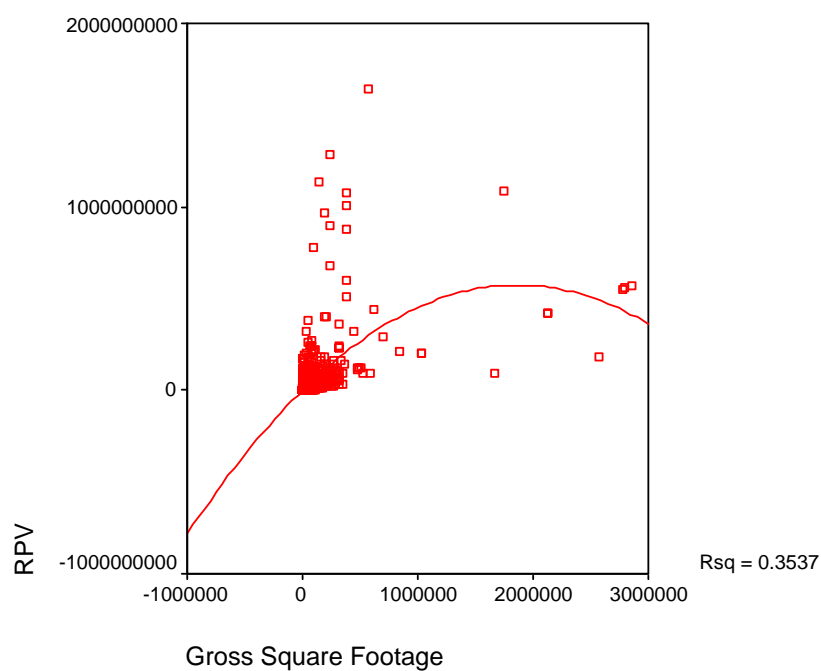


Figure 1-2-5. RPV vs. Gross Square Footage for All Facilities



We continued to discuss this issue, and options for examining it, with DOE during the early part of Phase II. For example, both sampling methods could have been applied at the pilot sites and their results compared.

In any case, we recommend including the 20 to 30 “outliers” identified in Figure 1-2-5 in the initial sample validation effort. The observations in the upper left of the chart (low square footage and high RPV) are likely to be “special” facilities, such as highly technical or complex small facilities, or serious data errors. Similarly, the observations in the lower right of the chart are likely to be large, simple facilities (for example, large parking lots) with very low replacement costs or serious data errors. Whatever their cause, these few obvious outliers should be evaluated as additional data points over and above the random sampled facilities. Outliers are a special case, shown as a definite source of error in our validation, and we include them to prevent skewed results.

FACILITY SELECTION

Although we address this topic in detail (see page 2-3-2), some discussion of the alternatives is warranted here. Either LMI or DOE headquarters can implement the sampling plan and selecting the actual facilities to be validated. A hybrid approach is also an option.

Briefly, the procedures for selecting a simple random sample would be as follows:⁹

1. Assign unique “case numbers” from 1 to n to each FIMS-reported facility in each program type at each installation in a spreadsheet.
2. Employ a computerized random number generator and apply it to the unique “case numbers” to select (a proportion of the total, on the basis of the sample size formula presented above) the actual facilities to be validated for each program at each site. Many statistical packages (for example, SAS or SPSS) can quickly and easily handle these two steps once the data are organized by program at each installation.
3. Print the list of selected facilities, by program, and provide it to each site along with the detailed validation procedures.

Due to the technical nature of the sample selection process, and because the data reside centrally in FIMS and the procedures are best carried out on a single PC at a single point in time against a single query of the FIMS database, we recommend not assigning this task to the individual installations; DOE headquarters or its agent should carry it out.

⁹ If the square footage weighting scheme is used, then an additional step (before 2), would be to add duplicate cases of each facility to the file in proportion to each facility’s gross square footage as a percentage of the total for that program at that site (or another appropriate weighting scheme, based on square footage, to achieve the same effect of weighting the probability of selection by square footage).

Chapter 2

Validation Standard

INFORMATION SOURCES

In developing a validation standard for the FIMS database, we considered the following sources of information:

- ◆ *FIMS database.* With assistance from OECM and OAI, we reviewed the scope and hierarchy of the database, values and ranges of key data elements, standard report outputs, and custom data queries.
- ◆ *FIMS website and official documentation.* We reviewed the FIMS users guide, training manual, data dictionary, and other related documentation provided by OECM.
- ◆ *OECM personnel.* We met with OECM to review in detail each metric and related data elements and to understand the overall uses of the FIMS database and underlying implications of accuracy.
- ◆ *Best practices.* We reviewed the data validation practices of several public and private-sector organizations to align our approach, where appropriate, with the best practices of those organizations. Organizations we looked at include NAVFAC, ETA, USACE, EPA, and RTI.¹

STATISTICAL APPROACH

To develop a statistical standard for validating the metrics, we considered the following:

- ◆ Their uses
- ◆ The level of error allowed or tolerated in the measurement of the metrics
- ◆ The desired level of confidence in the estimated accuracy of the metrics, which depends (in part) on their uses.

DOE primarily uses the metrics in budget formulation and budget execution related to facility repair and replacement. Although important, these uses are not urgent compared with, say, determining whether exposure to a given substance will cause a serious illness or disease. In the latter case, a level of allowable (tol-

¹ NAVFAC = Naval Facilities Engineering Command; ETA = Employment and Training Administration, USACE = U.S. Army Corps of Engineers, EPA = U.S. Environmental Protection Agency; and RTI = Research Triangle Institute.

erable) error might be 1 percent or less, and the desired level of confidence in the accuracy of the measurement might be 95 percent, or even 99 percent.

In addition, for the purposes of the validation process, it is only necessary to determine whether a site's and a program's reported metrics are "accurate or inaccurate within a level of tolerance," and not the "total estimated amount of inaccuracy" in the metrics. Given the budgetary use of the site- and program-level metrics, we recommended

- ◆ an allowable (tolerable) error of 10 percent, and
- ◆ a level of confidence of 90 percent that the estimated accuracy of the metric will be within the tolerable level.

As a result, we developed a "simple random sampling" plan that allows DOE to make the following type of statements at the end of the validation process:

- ◆ *Site level.* For a given program at a given site, we are 90 percent confident that x percent of asset RPVs are accurate within ± 10 percent. (A similar statement can be made for the ACI and AUI metrics for each program at each site.)
- ◆ *Program level.* For a given program overall, we are 90 percent confident that x percent of asset RPVs are accurate within ± 10 percent. (A similar statement can be made for the ACI and AUI metrics for each program overall.)

These statements are with respect to "facilities," not "square footage." In other words, under the simple random sampling plan, each facility in each program at a given site has an equally likely chance of selection for validation, regardless of size. It may be desirable to weight the sample selection process by a measure of square footage, giving larger facilities a higher chance of being included in the sample and ensuring representation across the size distribution of facilities.² In either case, the sample size for each separate program at each site varies between 1 and 25, depending on the total number of facilities in each program at the site. We use the same sample of each program at each site to determine the accuracy of each of the three metrics and to calculate the estimated accuracy of the metrics by program and site.

In Phase II, we develop a process for determining "root causes" of errors identified during the validation. This process focuses on determining, for each facility found to be "outside the allowable error of 10 percent" for a given metric, the probable sources of the error. We develop guidelines for using this information to conduct "root cause analysis" to identify the major sources of error and to aid in developing error-mitigation strategies.

² We recommend a method in the second part of this report.

CALCULATIONS

The approach that we recommend—determining whether the metric for a facility is accurate or inaccurate—results in a situation characterized in statistics as a “binomial distribution.” In other words, a portion, p , of the population has some attribute, and the remainder of the population, $1-p$, or q , does not.³ In this case, the attribute is “accuracy of the metric for a particular facility.”

We have no strong reason to suspect that the accuracy of any one metric significantly differs from the accuracy of another. In fact, each one contains common items (for example, square feet), so the accuracy of each metric for a given facility should be similar and related to the accuracy of the other metrics for that facility. For these reasons, our calculation of sample size for each program and site uses a single estimate of p (and therefore of q) equal to 0.9. In other words, we expect that the accuracy of each metric, for each program and at each site, is approximately 90 percent.

As discussed previously, the amount of “allowable error,” L , in a particular facility’s metric is 10 percent, therefore the level of confidence desired in the accuracy of a metric for a program at a site is 90 percent (“0.90 confidence probability”). Using the formulas for sampling from a binomial distribution,⁴ we find that these assumptions yield the following sample size calculations for each separate program at each site:

$$n = \frac{2.7 pq}{L^2}, \quad [1]$$

or, since we are dealing with relatively small populations per site, we use a finite population correction factor:

$$n' = \frac{n}{\left(\frac{1+n}{N} \right)}, \quad [2]$$

where N is the total population of facilities of that program type at the installation, and n exceeds 10 percent of N .

Application of these formulas results in sample sizes (n or n') that vary from 1 to 25 for a program at a given site, depending on the total (N) of facilities of that particular program type at that site. Table A-1 shows the required sample size for each program at a site, using the above formulas, for all values of N up to 243⁵ as well as expected sample sizes by program by site (see Appendix C). If a 95

³ George W. Snedecor and William G. Cochran, *Statistical Methods*, sixth edition (Ames, Iowa: The Iowa State University Press, 1967), p. 202.

⁴ See Note 3, pp. 516–518.

⁵ After which the same size remains constant at 25.

percent level of confidence is desired, the maximum sample size for each program at each site increases to 36, which would result in validating up to 44 percent more facilities in each program at each site than required by a 90 percent level of confidence. The 5 percent increase in level of confidence does not justify the increased time and effort in this situation on the basis of our initial finding regarding the quality of data for facilities considered non-life-threatening.

We can “roll up” the results to the DOE level for each program for each metric. We simply put all sampled facilities in a given program across all sites into a single group and use these observations to calculate the accuracy (frequency of error and non-error) of the metrics for that program at the DOE level.⁶

SQUARE FOOTAGE WEIGHTING

As discussed previously, the above statements and approach are with respect to “facilities,” not “square footage.” In other words, with simple random sampling, each facility in each program at a given site has an equally likely chance of being selected for validation. It may be desirable to weight the sample selection process by a measure of square footage, giving larger facilities a higher chance of being included in the sample at a future date.

We analyzed the size (square footage) distribution of facilities for the individual programs and for all programs at the two proposed pilot sites; the size is not normally distributed. Our analysis indicates that, for most programs, the vast majority (about 85 percent) of facilities are relatively small (less than 10,000 square feet). Figures 1-2-1 through 1-2-3 show frequency histograms of gross square footage for all programs and separately for the NNSA and EM programs (which account for two-thirds of DOE facilities in FIMS).⁷ Eight-five percent of the facilities are small and were not included in the weighting process.

Simple random sampling (each facility has an equal chance of being selected for validation) results in very few large facilities. This approach and outcome is desirable if the focus is on the “incidence” (frequency) of error in FIMS. It is also desirable if mission criticality of facilities is unrelated, or negatively correlated, with size of facilities. For these reasons, we recommend that this first effort to validate FIMS use the simple random sampling approach, which provides an accurate estimate of the incidence or frequency of error and causes of error in the relatively new FIMS program. It also allows for policies and actions to reduce the frequency of error across all facilities.

⁶ We are sampling to determine the accuracy of each program at each site and therefore have up to 25 sampled facilities in each program that has facilities at a given site. With respect to each program overall in DOE, this represents an “over-sampling” to achieve the required confidence as to the accuracy of each program in total.

⁷ Due to the wide range of values for gross square footage, we chose to display up to 285,000 square feet on the x-axis. The actual high range is about 2.8 million; if we displayed that scale, the graphs would be unreadable.

Figure 1-2-1. Frequency Distribution of Gross Square Footage for All Facilities

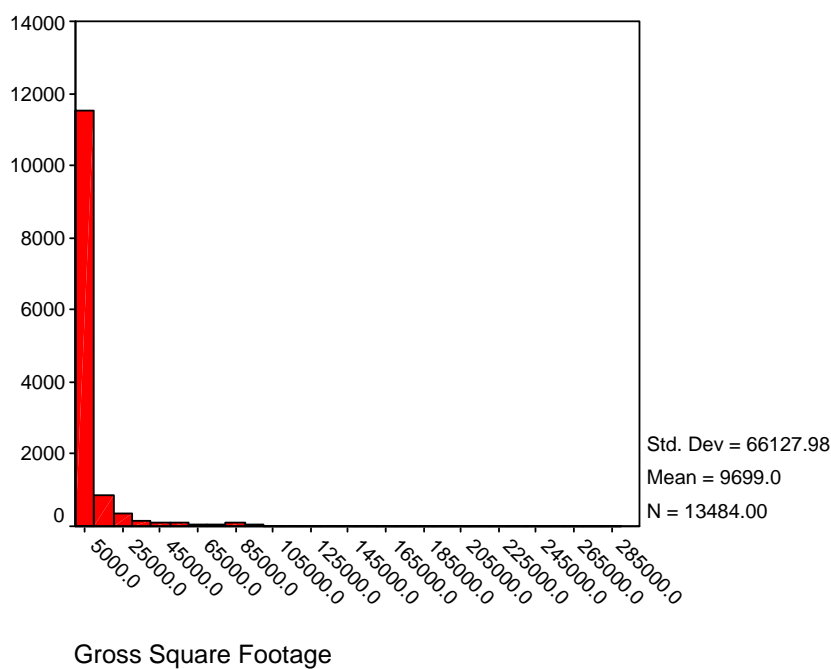


Figure 1-2-2. Frequency Distribution of Gross Square Footage for EM Program

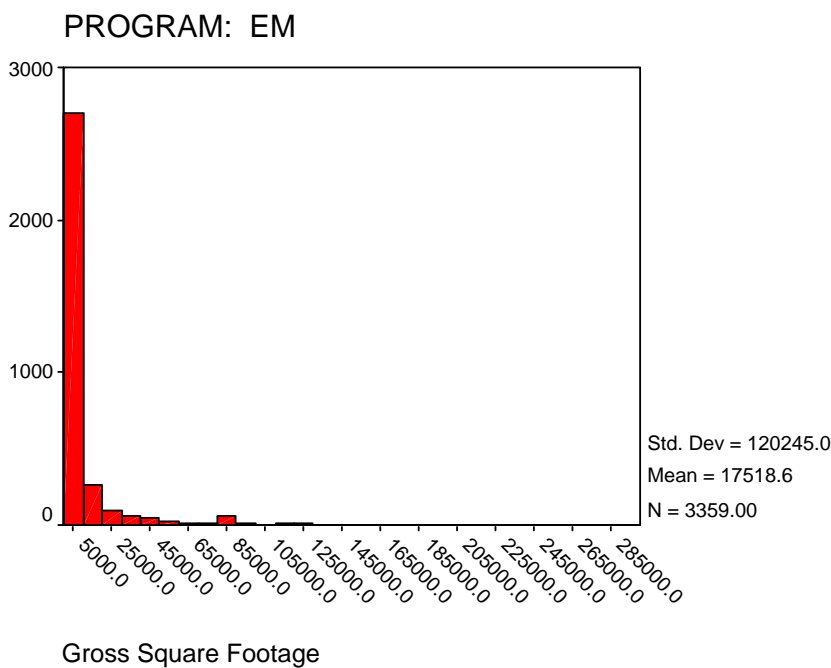
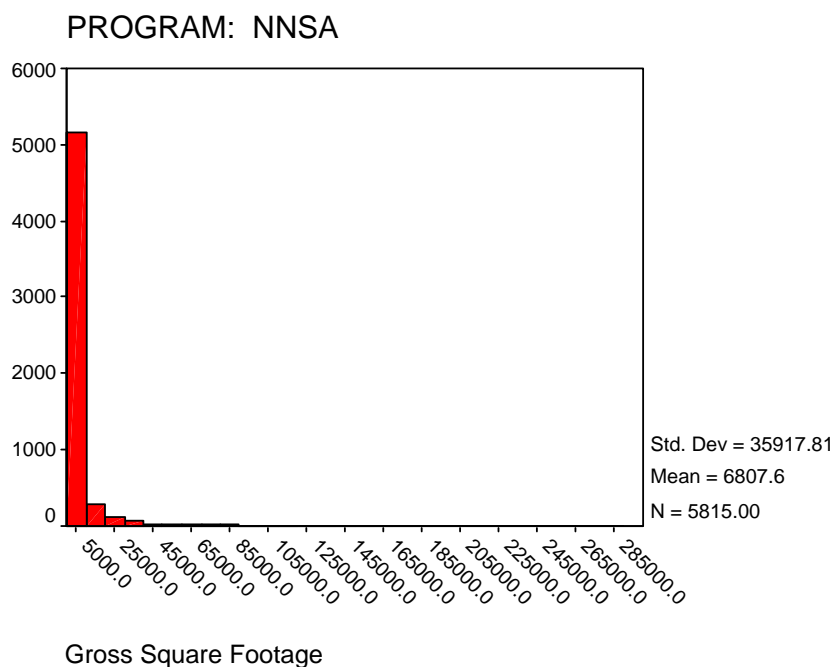


Figure 1-2-3. Frequency Distribution of Gross Square Footage for NNSA Program



Since square footage is an important driver of RPV, it is not surprising that about 80 percent of all replacement values are less than \$100,000, as can be seen from Figure 1-2-4 (a histogram of the frequency distribution of RPVs).⁸ The scatter-gram of RPV against gross square footage in Figure 1-2-5 shows that the overwhelming majority of DOE facilities are concentrated together in a range that can be characterized as “small with low replacement cost.” Together, these figures indicate that weighting the sample by square footage would not give different results than simple random sampling. The results would be the same regardless of their use. However, in future validation efforts (after correcting policies and procedures), we recommend that DOE use a weighted random sampling technique on one occasion, weighting each facility by its reported gross square footage as a percentage of total reported square footage for that program at that site. We do not expect a significant difference in results although the technique may provide a different estimate, compared with simple random sampling, of the accuracy of the metrics when used for budgetary purposes. This technique will allow the error rates of larger facilities to be compared with those of the overall population.

⁸ Again, for readability, we chose to display only a portion of the x-axis scale. The high end RPV range is about \$1.6 billion.

Figure 1-2-4. Frequency Distribution of RPV for All Facilities

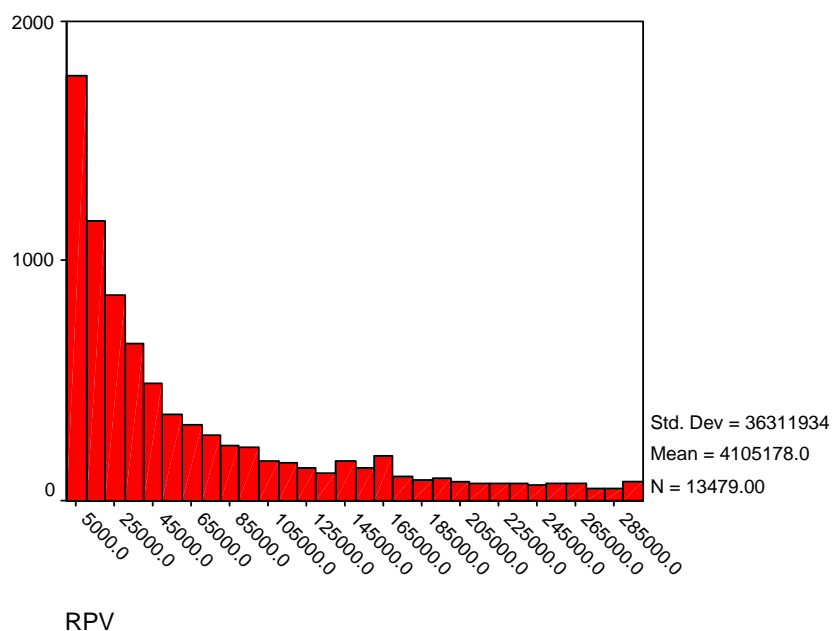
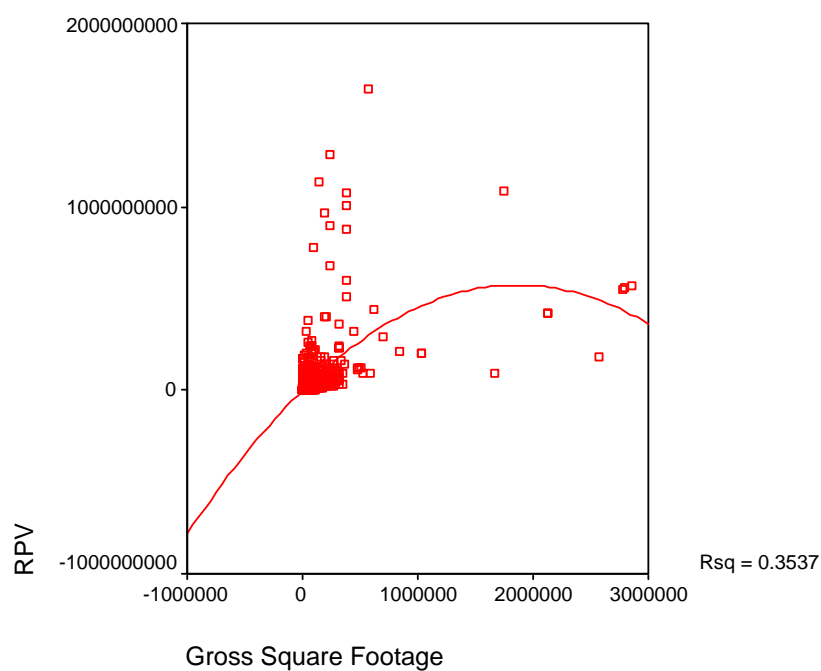


Figure 1-2-5. RPV vs. Gross Square Footage for All Facilities



We continued to discuss this issue, and options for examining it, with DOE during the early part of Phase II. For example, both sampling methods could have been applied at the pilot sites and their results compared.

In any case, we recommend including the 20 to 30 “outliers” identified in Figure 1-2-5 in the initial sample validation effort. The observations in the upper left of the chart (low square footage and high RPV) are likely to be “special” facilities, such as highly technical or complex small facilities, or serious data errors. Similarly, the observations in the lower right of the chart are likely to be large, simple facilities (for example, large parking lots) with very low replacement costs or serious data errors. Whatever their cause, these few obvious outliers should be evaluated as additional data points over and above the random sampled facilities. Outliers are a special case, shown as a definite source of error in our validation, and we include them to prevent skewed results.

FACILITY SELECTION

Although we address this topic in detail (see page 2-3-2), some discussion of the alternatives is warranted here. Either LMI or DOE headquarters can implement the sampling plan and selecting the actual facilities to be validated. A hybrid approach is also an option.

Briefly, the procedures for selecting a simple random sample would be as follows:⁹

1. Assign unique “case numbers” from 1 to n to each FIMS-reported facility in each program type at each installation in a spreadsheet.
2. Employ a computerized random number generator and apply it to the unique “case numbers” to select (a proportion of the total, on the basis of the sample size formula presented above) the actual facilities to be validated for each program at each site. Many statistical packages (for example, SAS or SPSS) can quickly and easily handle these two steps once the data are organized by program at each installation.
3. Print the list of selected facilities, by program, and provide it to each site along with the detailed validation procedures.

Due to the technical nature of the sample selection process, and because the data reside centrally in FIMS and the procedures are best carried out on a single PC at a single point in time against a single query of the FIMS database, we recommend not assigning this task to the individual installations; DOE headquarters or its agent should carry it out.

⁹ If the square footage weighting scheme is used, then an additional step (before 2), would be to add duplicate cases of each facility to the file in proportion to each facility’s gross square footage as a percentage of the total for that program at that site (or another appropriate weighting scheme, based on square footage, to achieve the same effect of weighting the probability of selection by square footage).

Chapter 3

FIMS Metrics

In this chapter, we discuss the three key FIMS metrics for which we develop a statistical standard for validating accuracy.

REPLACEMENT PLANT VALUE

RPV is the Department's corporate measure of the cost to replace the existing structure with a new structure of comparable size using current technology, codes, standards, and materials.

FIMS uses RPV models¹ to provide standard, justifiable building costs for the DOE building inventory. RSMeans annually updates the RPV models on the basis of private-sector construction costs, adjusted to a nationwide average per square foot.² This is the starting basis for developing the RPV value. DOE adjusts this national cost to include differences in geographic location (also provided by RSMeans annually) and inputs it into the FIMS system. The geographic factor is used to adjust the wage rates and material costs for the local area.

In addition to the geographic location factor, DOE applies a site factor to the RPV, to adjust for costs for security, site fees, permitting fees, construction management services, preparation of as-built drawings, startup, and commissioning fees specific to the site. The FIMS default site factor is 1.568, but each site has a process for determining its site factor. The site factor varies from building to building on the basis of type, size, and facility makeup. The addition of geographic and site factors results in a more accurate total construction budget cost for the building. The adjusted RPV unit cost is multiplied by the gross square footage of the building to determine the final RPV cost:

$$RPV = gross\ sq\ ft \times RPV\ unit\ price \times geographic\ cost\ factor \times site\ factor.$$

DOE values for the RPV range widely due to the vastly different types and sizes of facilities that span the asset inventory. RPVs range from just over \$1,000 to over \$1 billion. Roughly 50 percent of all RPVs are \$100,000 or less, and 99 percent are \$100 million or less. The remaining 1 percent, however, range from \$100 million to \$1.6 billion. Although the RPV adheres to a relatively normal distribution, the gross square footage, a key driver of RPV, is skewed toward the lower end of the value range. Gross square footage values range from under 10 square

¹ RSMeans provides the RPV values.

² RSMeans, a leading supplier of construction cost information, offers annual construction cost data books, construction estimating and facilities management seminars, electronic cost databases and software, reference books, and consulting services.

feet to over 2.8 million, but 85 percent of facilities are at 10,000 gross square feet or less.

Because the square footage and site factors are the RPV formula items that the site inputs, factors affecting the estimate of these two items are the most likely sources of error. Many sites calculate their own RPV using general engineering estimates or other methods that are not always consistent, and others wait “too long” to update a construction cost (a situation in which the RSMeans model is not used). DOE permits the sites to estimate and input their own RPV numbers in place of the FIMS models. The RS Means-derived RPV units are too generic to represent a large portion of the buildings. Many of the values are too low, and the sites have altered their values to more accurately represent the true value of their buildings.

ASSET CONDITION INDEX

The ACI is the department’s corporate measure of the condition of its facility assets. It reflects the outcomes of real property maintenance and recapitalization policy, planning, and resource decisions. The ACI is calculated as

$$ACI = 1 - FCI,$$

where the facility condition index (FCI), is the ratio of deferred maintenance to RPV. The FCI is derived from data in FIMS.

Ratings are assigned to ACI range measures. The goal is for the ACI to approach 1. The ACI increases and approaches 1 as the condition of facilities improves at a site, that is, deferred maintenance approaches zero. ACI ranges and ratings are as follows:

- ◆ 1.00 to 0.98, Excellent
- ◆ 0.98 to 0.95, Good
- ◆ 0.95 to 0.90, Adequate
- ◆ 0.90 to 0.75, Fair
- ◆ Less than 0.75, Poor.

The desired quality and accuracy of each metric are the same (see page 1-2-3): a 10 percent variance tolerance in the calculation of the index and a 90 percent level of confidence in the estimated aggregated metric for a site or program. Therefore, the sample size for determining the accuracy of the ACI by program within each site and overall is the same as the sample size requirement for the RPV. Consequently, the same sample is used to estimate the accuracy of each index, for each program at each site.

About 90 percent of the ACI values we saw in the database were greater than 0.75, 8 percent were less than 0.75, and about 2 percent were negative. A negative ACI value indicates that the deferred maintenance value is higher than the RPV. In one extreme case, the deferred maintenance was listed as \$122,000, while the RPV for that facility was only \$6,000. The definition of deferred maintenance—project cost, deficiency cost, or total replacement cost—varies from site to site; the department has yet to clearly define what should be captured as deferred maintenance amounts. We need to better understand the validity of the last 10 percent, particularly the negative values, before we can develop a meaningful validation process for the ACI metric (see page 2-5-4).

Because the ACI is a function of the FCI, its accuracy is a function of the elements of the FCI, especially those that are not automatically determined by the RSMeans model. These elements are the condition assessment code, the amount of deferred maintenance, and square footage. An additional important source of error is the length of time elapsed since the last condition assessment (up to 5 years, depending on mission criticality). Also, for the ACI, in the validation process, we include procedures for comparing critical variables that likely depend on (correlate with) one another. Examples of likely dependency, or correlation, include the age of a facility compared with its condition assessment and the date it was last recapitalized.

DOE has not mandated one process for condition assessment surveys. Each site is allowed to use its own process to cost-estimate deficiencies as long as the process involves a “nationally recognized product.”

ASSET UTILIZATION INDEX

The AUI is the Department’s corporate measure of facilities and land holdings compared against requirements. The index reflects the outcome from real property acquisition and disposal policy, planning, and resource decisions. The index is the ratio of the area of operating facilities or land holdings justified through annual utilization surveys (numerator) to the area of all operational and excess facilities or land holdings without a funded disposition plan (denominator). The AUI is derived from data in FIMS obtained from annual utilization surveys.³ Separate AUIs are developed for facilities and land holdings.

Both numerator and denominator are measured in square feet:

$$AUI = \frac{\text{assets with justified utilization}}{\text{all current real property assets}}.$$

³ DOE O 430.1B, U.S. Department of Energy, page 15. See also 41 CFR, Chapters 101 and 102, and Executive Order 12512, Federal Real Property Management Survey.

Ratings are assigned to AUI range measures. The AUI improves as excess facilities are eliminated and consolidation increases the space utilization rate of the remaining facilities. AUI ranges and ratings are as follows:

- ◆ 1.00 to 0.98, Excellent
- ◆ 0.98 to 0.95, Good
- ◆ 0.95 to 0.90, Adequate
- ◆ 0.90 to 0.75, Fair
- ◆ Less than 0.75, Poor.

AUI and ACI measure the net result of numerous real property management and disposal policy, planning, and resource decisions over time. The Planning, Programming, Budgeting, and Evaluation cycle (PPBE) requires accounting for execution of resource decisions made during planning, programming, and budgeting. To assess the use of real property asset budgets for their intended purposes, DOE has the following execution measures:

- ◆ Headquarters compares the sites' real property maintenance and disposition budget execution with the amounts shown in the Integrated Facilities and Infrastructure (IFI) crosscut budget
- ◆ Headquarters program offices submit assessments of IFI crosscut budget execution for real property maintenance and disposition to OECM within 45 calendar days of the end of each fiscal year quarter.

Again, the desired quality and accuracy of each metric is the same: a 10 percent error tolerance in the calculation of the index, and a 90 percent level of confidence in the estimated accuracy of the index. Therefore, the sample size for determining the accuracy of the AUI by program within each site and overall is the same as the sample size requirement for the RPV and ACI. Consequently, the same sample is used to estimate the accuracy of each index, for each program at each site.

The components of the AUI are status utilization and net occupiable square feet. Both measures are potential sources of error, but especially the status utilization element. In addition, the status utilization can be miscoded—by mistake, because it has not been updated in FIMS for a long time, or because it has been coded with an inaccurate usage (building status code).

Chapter 4

Conclusions and Recommendations

On the basis of our analysis of the three key metrics, RPV, ACI, and AUI, in the FIMS database and in developing our sampling approach, we recommend the following.

Validate all three metrics using the same sample.

Since the components of each metric are related, and often exist in multiple indexes, the same sample can be used to validate all three metrics. The three metrics interrelate on the basis of common data elements within their respective formulas. Gross square feet (GSF) is a key data element in calculating RPV and AUI, and potentially a large source of error. Since RPV is used to calculate ACI, GSF also is a key element of that calculation.

Use a 90 percent level of confidence for statistical sampling and analysis.

A minimum site sample size of up to 25 is required to achieve at least 90 percent level of confidence.¹ A minimum site sample size of up to 36 would be required to achieve a 95 percent level of confidence—a 44 percent increase in sampling size and likely a commensurate level of required effort. For the purposes of budgetary decisions, the achievable additional benefit does not warrant the additional level of effort.

Use simple random sampling at first.

About 85 percent of all facilities are less than 10,000 square feet and about 50 percent have replacement costs less than \$100,000. Therefore, simple random sampling should be used. Future validations can weight sampling by facility square footage and then compare the results with those of simple random sampling.

Sample all outliers.

The 20 to 30 outliers (Figure 1-2-5), defined as high gross square feet with low RPV or low gross square feet with high RPV, should be added to the randomly sampled facilities in order to understand whether data for these facilities are accurate.

¹ See Appendix C-1, table C-1 for list of required sample size by site population.

Focus the validation on the incidence or frequency of errors in each metric and the main sources or causes of those errors.

DOE does not use a mission criticality index to segregate or prioritize the facilities and their metrics. Focusing on the frequency of errors first, providing the most objective learning process in understanding the main sources of error, best serves the validation.

Sample each program separately at each site.

Several sites contain facilities belonging to more than one program. To maintain our ability to roll up data at the program level, we should treat each program within a site as a separate site. This results in samples of up to 25 facilities of each program at a given site. When the site is managed by multiple programs, take one sample.

FACILITIES INFORMATION MANAGEMENT SYSTEM

PHASE II. METRICS VALIDATION PROCESS

Odoi Associates Inc.

Robert E. Ward

Brent L. Vanderpool

LMI

Richard A. Skulte

Jonathan P. Adams

Robert L. Crosslin

Anna M. Wallace

Gerald W. Westerbeck

SEPTEMBER 2004

Chapter 1

Introduction

The Department of Energy’s Office of Engineering and Construction Management asked LMI—through Odoi Associates, Inc. (OAI)—to develop a validation process for its Facilities Information Management System (FIMS) data. This part of our report, Phase II, describes the details, definitions, and logical sequence of the validation process, which we used in pilot validations at two DOE sites. LMI evaluated the results of the pilot, documented findings and observations, and incorporated feedback and lessons learned.

OBJECTIVE

The objective of Phase II was to develop a process for use in validating the accuracy of three key FIMS metrics at the site and program levels:

- ◆ Replacement plant value (RPV)
- ◆ Asset condition index (ACI)
- ◆ Asset utilization index (AUI).

As part of this objective, we recommend a logical sequence for checking FIMS data elements against their source documents and reporting the results.

STRUCTURE

The remainder of this part of our report comprises the following chapters:

- ◆ In Chapter 2, we describe the pilot validations, setting expectations, and providing general guidance.
- ◆ In Chapter 3, we detail the draft validation process, describing each logical step in sequence.
- ◆ In Chapter 4, we present our findings and observations on the pilot process and recommend adjustments to the validation standard and process.
- ◆ In Chapter 5, we draw conclusions and recommendations for full implementation of the validation process.

Chapter 2

Pilot Validation

PURPOSE

The pilot allowed us to test and refine the approach and procedures for validating FIMS. The numerous sites and facilities vary in size and usage, so learning as much as possible during the pilot phase was critical to ensuring that the full validation process can be carried out efficiently and effectively to meet DOE's objectives.

EXPECTATIONS

We executed the pilot test in two sites—the National Energy Technology Laboratory and the Brookhaven National Laboratory—which are somewhat dissimilar. The lessons we learned at these pilot sites helped us revise the final procedures and templates:

- ◆ *Clarity of procedures.* In statistical estimating processes, procedures need to be followed consistently by all participants and applied consistently to all sites and facilities. The pilot helped us clarify the operating procedures for validating each data element.
- ◆ *Clarity of informational requests.* The sources for validation and the sources for material variation will help DOE determine needs for standardization and corrective action programs. The pilot helped us to clarify the information requests for documenting data sources and reasons for material variations between sources and FIMS.
- ◆ *Ease-of-use of reporting format.* All sites will use the FIMS validation and reporting template repeatedly for years to come. We needed to learn as much as possible about this template before programming the online applications that all sites will use.
- ◆ *Types of source documents for each data element.* FIMS is a quickly accessible central repository of critical information on all DOE facilities. However, the detailed comprehensive data that should be the “accurate data of record” about each facility exist in source documents at the sites. The pilot told us about the nature and frequency of the various types of source documents that can be expected during the full validation and will allow for a more thorough set of examples to be included in the final users guide.

-
- ◆ *Sources of variation.* Sources of variation between source documents and FIMS are critical in understanding the reasons for inaccuracies and developing mitigation strategies. The final database application tool should include a “pick list,” from which users can choose, of the many potential sources of variation. The pilot provided this information.
 - ◆ *Time and effort required for validation.* The validation process should take a reasonably short time and consume a reasonable amount of staff effort: the time and effort required should be offset by the value of the results. The pilot provided information on lengthy and unnecessary steps, and on organizing the procedures to efficiently use the validation team’s time.
 - ◆ *Appropriate makeup of validation teams.* The pilot helped determine the type and amount of resources needed for validation at a site, including the value of a site-independent team member.

During the validations at the pilot sites, headquarters and LMI team members took extensive notes, including comments from site team members, regarding all aspects of the validation process. In addition, we conducted out-briefs to discuss the results of the validation, as well as lessons learned, with team members at each site. We recommended a 2-day validation at each pilot site.

The process is designed as a “desk” validation. We do not require the sites to measure square footage, tour the facilities to assess conditions, or perform other detail-level activities. Some source documents used for the validation for important data elements may be out of date (more than 3 years old) or otherwise incorrect, resulting in the validation of a significant number of the sampled facilities metrics as “correct” against these incorrect sources (in other words, a “false positive” finding). Alternatively, metrics may be deemed inaccurate due to outdated source documents, but still be correct (a “false negative” finding).

We needed to determine the extent of these possibilities, and their potential impact, during the pilot phase. Therefore, we asked the pilot sites to “sample the sample.” LMI randomly identified a limited number of the sampled facilities for detail-level validation. This provided us with the relative accuracy, and therefore the reliability, of the desk validation approach.

We do not assume that the pilot sites represent either a program or DOE as a whole. For these purposes, we will enter the individual sites during the full validation into a single database, from which we will calculate aggregate statistics for each metric and data element. For a program, we will calculate the frequency of error or material variation of FIMS from source documents as the total number of errors for a metric divided by the total number of sampled facilities for that program. We will include, for DOE in total, all facilities for all programs.

Chapter 3

Validation Process

OVERVIEW

In this chapter, we detail the validation process. Since the three metrics and several data elements are interrelated, we present the validation process in a logical sequence so that, for example, the validation of the first metric, RPV, can feed into the validation process of the second metric, ACI. Our goal in sequencing was to eliminate redundancy from the validation process.

The validation process consists of three phases: *initialization*, *validation*, and *evaluation and reporting*. In the following paragraphs, we work through each phase in detail.

INITIALIZATION

Assemble Validation Team

The ideal validation team consists of three people. Two should be intimately familiar with the facilities to be validated, their history, and source documents. The third should be a neutral party, someone not associated with the program or site being validated. This team member can be a representative from another DOE program, someone from the site's service center, someone from headquarters, or a designated third party.

Initially, DOE may wish to dedicate a small committee to participate as the neutral third member of the validation team, going from site to site, when possible, for the first wave of validations. This observation process continues the learning process of the pilot studies and more readily reveals the validation issues common to multiple sites.

Determine Validation Schedule

To statistically validate the accuracy of FIMS for each major program and DOE-wide, each site must be sampled. This also allows accurate validation for the individual sites. Scheduling of individual sites can be a matter of convenience and efficiency for the validation teams. The timing of individual sites does not affect the aggregate accuracy of either the program or DOE. However, a program's accuracy is not validated until all sites in the program are complete, and DOE's accuracy is not validated until all sites are complete.

Several weeks before validating an individual site, the validation team should send a copy of the blank FIMS validation and reporting template to the site, as well as a proposed agenda or schedule. The team should also inform the site that it needs access to drawings (AutoCAD, for example), work requests, work orders, FIMS, other databases that track site facilities, Condition Assessment Information System (CAIS), and people who use and maintain these records or systems day to day.

Generate Sample Set

We anticipate a sampling process that is an online application attached to the web version of FIMS. A short time (2 days, for example, because some FIMS values can change often) before the scheduled start date for a site's validation, headquarters (a dedicated FIMS validation person or team) uses the application to generate the sample facilities for that site. The sampling application performs the following tasks:

- ◆ Identify all facilities at the site by labeling each building with an identification number, such as MARS #001, MARS #002, and MARS #003.
- ◆ Apply filters to identify the subset of facilities to be sampled, such as buildings and real property trailers (MARS # 001), no leased facilities, no personal property trailers (MARS # 025) and no facilities identified for demolition or excess.
- ◆ Access the sample size lookup table (Appendix A) to determine the required sample size for the site.
- ◆ Use a random number generator to obtain a list of random numbers equal in quantity to the required sample size and associate those random numbers with specific facilities at the site.
- ◆ Identify “outliers” (from parameters set by OECM) to be added to the random sample. Examples of outliers include facilities with exceptionally high RPV and low gross square footage or exceptionally high gross square footage and low RPV (Figure 1-2-5). We include these outliers in the sample initially to better understand their unique qualities.
- ◆ Download the FIMS values for the metrics and associated data elements for the sample-selected facilities (plus any outliers) to the FIMS validation and reporting template (see Appendix B) for that site.

Headquarters then notifies the site and its validation team that the FIMS validation and reporting template is available online for its sampled facilities. This entire process can be automated such that headquarters can select, automatically generate, and notify all sites, a portion (such as a program), or an individual site.

Identify Source Documents

The site validation team uses the FIMS validation and reporting template as its basic tool for the validation process. FIMS values should be backed up by source documents, and methods surrounding those source documents, that could support an audit. Therefore, the first step is to gather available source documents for each required data element. Possible sources include

- ◆ construction drawings and blueprints,
- ◆ as-built drawings,
- ◆ condition assessment reports,
- ◆ work order requests and orders,
- ◆ other DOE databases (CAIS or Maximo only for FIMS validation),
- ◆ DOE-specific models and benchmarks (RSMeans, BOMA, etc.),
- ◆ physical measurements, and
- ◆ internal memorandums or other documented reporting.

Most or all of the sites maintain other (site- or program-specific) facility databases, which may contain some or all of the FIMS data to be validated. These databases, like FIMS, are repositories of information and may not be the true source documents for the purposes of FIMS validation—this needs to be understood by headquarters and the site validation teams. As these and additional sources are identified, we recommend tracking and tabulating them during the validation process—the FIMS validation and reporting template contains fields, or space, for this purpose. In doing so, DOE can identify trends and document needs for possible standardization in the future.

Baseline FIMS Data for Validation at Site

Before each validation, DOE headquarters prepopulates the FIMS validation and reporting template with data from FIMS by generating the random sample of facilities to be validated. This data set includes the facility name and identification number and has a date stamp. These data serve as the baseline data to be validated. To ensure integrity in the validation process, the site validation team must use the data recorded in FIMS, and only this, date stamp and not data recorded in FIMS or any other facilities databases after this date.

VALIDATION

Validation Reporting Form

For each required data element (see Tables 2-3-1 to 2-3-3), the validation team enters the following into the FIMS validation and reporting template:

- ◆ *Source*. The most recent source document for the data element in question.
- ◆ *Year of source*. The year the source document was created, modified, or last updated. If the source document is older than 1 year for utilization or 3 years for DM or CAS inspection if mission essential, 5 years if not mission essential, 1 year for RPV, or 1 year for site factor (from the date of the validation), then the value of that FIMS metric is “automatically in error (a material variance).”
- ◆ *Value from source*. The value from the source document that corresponds to the data element in question. In the case of square footage, it may be necessary to extract individual measurements from drawings and calculate a total measurement.
- ◆ *Most likely cause of variance*. The validation reporting form calculates the difference between the FIMS and source values to determine whether it is greater than 10 percent. If it is, the validation team should try to identify the likely cause of the variance. Possible causes of variance include
 - out-of-date (see Year of source above), inaccurate, or incomplete sources;
 - sources that probably exist, but cannot be located;
 - no sources;
 - multiple or conflicting sources;
 - incorrectly calculated values; and
 - incorrectly entered values.
- ◆ *Additional comments*. The validation team can use this space to document other observations or issues regarding the validation of FIMS data for a particular building or trailer or further explain data sources and variances.

Finally, the FIMS validation and reporting template automatically calculates the value of each of the three metrics using the formulas defined in the first part of this report, Phase I. It also compares this calculated value with the FIMS baseline values as well as with the value entered from the referenced data source. This ex-

ercise notes differences between hard-coded FIMS entries for the three metrics and their calculated values. Again, a variance greater than 10 percent raises a flag for follow-up and, if frequent enough, could indicate the need for possible standardization in the future. The FIMS validation and reporting template automatically calculates the frequency of error (in percentage points) for each of the three metrics, and each data element for each metric, for the entire site.

Required Data Elements

The metrics and associated data elements listed in the tables in this section are included in the validation. Some data elements factor into more than one metric, so the validation reporting form also reflects the sequence of data elements in the tables, which is the sequence we recommend for performing the validation.

REPLACEMENT PLANT VALUE

Table 2-3-1 shows the required data elements for the RPV metric validation.

Table 2-3-1. RPV Data Elements

Textual name	FIMS element name
Building RPV	PBLD_BUILDING_RPV
Gross square feet	PBLD_GROSS_SQFT
Site factor	PBLD_LAB_USAGE_PERCENT
Geographic cost factor	SITE_GEOCOST_FACTOR
RPV model type	RPVM_MODEL
RPV unit cost	RPVM_UNIT_COST
Adjustment cost	CAPI_IMPROV_COST
Adjustment date	CAPI_IMPROV_DATE
Adjustment description	CAPI_IMPROV_DESC

RPV is the department's corporate measure of the cost to replace the existing structure with a new one of comparable size using current technology, codes, standards, and materials. FIMS automatically calculates the RPV using model square foot cost, gross square footage, a geographic adjuster, and a local site factor. The resulting RPV is intended for macro analysis and not as a substitute for a detailed cost estimate. Each site has the option to replace a FIMS-generated RPV with a site-derived or engineered value:¹

$$RPV = \text{gross sq. ft} \times \text{RPV unit cost} \times \text{geographic cost factor} \times \text{site factor}.$$

¹ Department of Energy, *Facilities Information Management System (FIMS) User's Guide*, "Data Dictionary," August 2003, p. A-7. Available from http://fims.hr.doe.gov/Downloads/USER_GDE.PDF.

While an important metric itself, RPV is also a key component of the ACI metric. For this reason, RPV should be examined before proceeding to ACI analysis.

The following subsections describe the elements to be validated. Due to the complexity and potential for more sources of variance within RPV, we've added several secondary data checks involving nonprimary elements. Specifically, we check the adjustment date, cost, and description for additional verification of the timeliness and accuracy of the primary data elements as defined by the RPV formula.

Building RPV

What to check	RPV listed in FIMS for the building
Element (tab name)	PBLD_BUILDING_RPV (RPV tab—system generated)
How to check	From source documents
Why checked	It is a primary derived value with which field data are compared. It is the high-level, system-generated metric that is being validated. The final calculation of validated data should equal this value for the building being examined.

Gross Square Footage

What to check (first)	Gross square footage of the sample building
Element (tab name)	PBLD_GROSS_SQFT (Building dimensions tab)
How to check	From source documents
Why checked	It is a primary component of the RPV calculation for any building.
What to check (second)	The actual gross square footage of the sample building. This is the total floor area, exterior wall to exterior wall, of a building in square feet. For a single-story facility, this likely matches the building "footprint." For a multiple-story facility, the sum of the gross square feet for all floors equals the building gross square feet.
How to check	There are several means of checking gross square footage: <ol style="list-style-type: none">1. Desk analysis—Using BOMA measurements previously provided, determine and record the gross square feet of the sample building.2. Desk analysis—Using engineering drawings, including any as-built drawings, that give the dimensions of the building.3. Physical survey—Physically measure the dimensions of the building, exterior wall to exterior wall for each floor. Although it can be quite time-consuming, this method, if done correctly, provides the most accurate data.
Why it is checked	It is a primary component of the RPV calculation for any building.

Site Factor

What to check (first)	The site factor listed for the building
Element (tab name)	PBLD_LAB_USAGE_PERCENT (RPV tab)

How to check	From source documents
Why checked	It is a primary component of the RPV calculation for any building. This cost factor may not be the same for all buildings at any given site, given that different site considerations (security, environment, etc.) may exist within a single site and result in multiple site factors.
What to check (second)	Is the site factor listed for the building correctly determined for facilities of this type on this specific real estate at this site (check two or three others)?
How to check	From source documents
Why it is checked	It is a primary element of the RPV calculation.

Geographic Cost Factor

This factor is multiplied by the building or trailer RPV to adjust for local variations at DOE sites. It is for labor and material only and does not account for special site-related escalators.²

What to check (first)	The geographic cost factor listed for the building
Element (tab name)	SITE_GEOCOST_FACTOR (RPV tab)
How to check	From source documents
Why checked	It is a primary component of the RPV calculation for any building. This cost factor should be the same for all buildings at any given site.
What to check (second)	Is the geographic cost factor listed for the building the same as the geographic cost factor for other buildings at this site (check two or three others)?
How to check	From source documents
Why checked	If different, is there a reasonable explanation? (List it in the notes section of the checklist.)

RPV Model Type and Unit Cost

What to check (first)	RPV Model , a typical building that would be built to replace the existing sample building. ³ It uses costs and engineering statistics compiled by RS Means. Alternatively, if the RPV Model does not apply to this facility, check an appropriate source document.
Element (tab name)	RPVM_MODEL PBL_BUILDING_GROUP_NO (Lookup table, RPV tab)
How to check	From source documents
Why checked	As a secondary data element, the RPV model determines what RPV unit cost is used.
What to check (second)	RPV Unit Cost , a national unit cost for the RPV Model building. ⁴ It is calculated by dividing the total cost of the model by the square footage of the model.
How to check	From source documents

² See Note 1, p. A-16.

³ See Note 1, p. A-37.

⁴ See Note 1, p. A-38.

Element (tab name)	RPVM_UNIT_COST (Lookup table)
Why checked	It is a primary data element needed to determine RPV.

RPV Data Entry—Secondary Checks

What to check (first)	The date of acquisition /construction (age) of building
How to check	Interview site facility managers
Why checked	For new facilities less than 5 years old, the actual acquisition (construction) cost is the best indicator of RPV.
What to check (second)	For new buildings, the cost of acquisition /construction (age) of building
How to check	Interview site facility managers
Why checked	For new facilities less than 5 years old, the actual acquisition (construction) cost is the best indicator of RPV.
What to check (third)	Adjustment Cost ⁵
Element (tab name)	CAPI_IMPROV_COST (Cap adjust tab)
How to check	From source documents
Why checked	A secondary data element, this data entry is made by site accounting or finance to indicate the cost of a capital adjustment or improvement. It requires an engineering assessment of the adjustment description to determine whether the adjustment materially increased the RPV of the sample building. In many instances, capitalized improvements increase the RPV of a building if the improvement increases the replacement cost of the building. Such improvements might include expansion of the building or conversion to another use—such as converting storage to administrative space. However, other capital improvements, such as relocating walls in an administrative building, might not materially increase the RPV of a building. Thus, an assessment is required if an adjustment has been made to the building.
What to check (fourth)	Adjustment Date ⁶
Element (tab name)	CAPI_IMPROV_DATE (Cap adjust tab)
How to check	From source documents
Why checked	A secondary data element, it is needed to determine the relative accuracy of the adjustment cost data. The older the adjustment is, the less valid. (Note any adjustments older than 5 years.)
What to check (fifth)	Adjustment Description ⁷
Element (tab name)	CAPI_IMPROV_DESC (Cap adjust tab)
How to check	From source documents
Why checked	A secondary data element, this data entry is made by site accounting or finance to describe a capital adjustment or improvement. It requires an engineering assessment of the adjustment description as to whether the adjustment materially increased the RPV of the sample building by the amount indicated in the adjustment cost entry. In many instances, capitalized improvements increase the RPV of a building if the improvement increases the replacement

⁵ See Note 1, p. A-3.

⁶ See Note 1, p. A-3.

⁷ See Note 1, p. A-3.

cost. Such improvements might include expansion of the building or conversion to another use—such as converting storage to administrative space. However, other capital improvements, such as relocating walls in an administrative building, might not materially increase the RPV. Thus, an assessment is required if an adjustment has been made to the building.

ASSET CONDITION INDEX

Table 2-3-2 shows the required data elements for the ACI metric validation.

Table 2-3-2. ACI Data Elements

Building ACI	FIMS report
Building ACI	FIMS report
Building FCI	FIMS report
Deferred maintenance	DEFM_DEF_MAINT
Building RPV ^a	PBLD_BUILDING_RPV

^a The results obtained in the RPV metric validation are used as input to the ACI metric validation.

The ACI is the department's corporate measure of the condition of its facility assets. It reflects the outcomes of real property maintenance and recapitalization policy, planning, and resource decisions. The ACI is calculated as

$$ACI = 1 - \frac{DM}{RPV}, \text{ or}$$

$$ACI = 1 - FCI,$$

where,

ACI = asset condition index,

DM = deferred maintenance,

RPV = replacement plant value, and

FCI = facility condition index (ratio of *DM* to *RPV*).

The goal is for ACI to approach 1. As DM is reduced, FCI decreases and ACI increases. ACI approaches 1 as the conditions of the site facilities improve.⁸

Because RPV is one of the three key metrics, as well as a key component in determining ACI, RPV should be determined as the first element in FIMS validation.

⁸ See Note 1, p. A-5.

Asset Condition Index

What to check (first)	The ACI listed for the sample building
Element (tab name)	Report generated
How to check	From source documents
Why checked	If not listed or different from the value separately calculated, is there a reasonable explanation? (Explain in the notes section of the checklist.)

Facility Condition Index

What to check (first)	The FCI listed for the sample building
Element (tab name)	Report generated
How to check	From source documents
Why checked	The FCI is the ratio of deferred maintenance (the cost of deficiencies of facility assets) to the facility's RPV. ⁹
What to check (second)	The Deferred Maintenance (DM) listed for the sample building
Element (tab name)	DEFM_DEF_MAINT (Building/trailer/OSF maintenance tab)
How to check	From source documents
Why checked	The deferred maintenance represents the maintenance deficiencies of a facility or site. It is maintenance that was not performed when it should have been or when scheduled and which, therefore, is put off or delayed to the future. ¹⁰ It is a key component of determining the ACI.
What to check (third)	Is the DM total for the building correct?
How to check	From source documents
Why checked	If different, is there a reasonable explanation? (Explain in the notes section of the checklist.)

ASSET UTILIZATION INDEX

Table 2-3-3 shows required data elements for the AUI metric validation.

Table 2-3-3. AUI Data Elements

Building AUI	FIMS report
Asset utilization index	FIMS report
Net occupiable square footage	PBLD_NET_OCC_SQFT
Gross square feet	PBLD_GROSS_SQFT
Building status	PBLD_CMST_STATUS
Status utilization	PBLD_PERCENT_UTILIZATION
Status code date	PBLD_CMST_STATUS
Excess indicator	PROP_EXCESS_IND
Excess year	PROP_EXCESS_YR

⁹ See Note 1, p. A-15.

¹⁰ See Note 1, p. A10.

The AUI is the department's corporate measure of facilities and land holdings compared with requirements. The index reflects the outcome from real property acquisition and disposal policy, planning, and resource decisions. The index is the ratio of the area of operating facilities or land holdings justified through such means as annual utilization surveys (numerator) to the area of all facilities (operational and excess) or land holdings without a funded disposition plan (denominator). The AUI is derived from data in FIMS obtained from annual utilization surveys.¹¹ Separate AUIs are developed for facilities and land holdings. Both numerator and denominator are measured in net occupiable square feet:

$$AUI = \frac{\text{assets with justified utilization}}{\text{all current real property assets}}, \text{ or}$$

$$AUI = \frac{\text{operational net occupiable square feet} \times \text{status utilization}}{\text{total net occupiable square feet (not in D\&D)}}.$$

AUI is generally a site or programmatic metric, providing information for an aggregate of facilities. The inclusion of square foot measurements in both the numerator and denominator ensures a proper weighting of large and small facilities. To truly reflect asset utilization of a site or program, all operating facilities need to be included. The FIMS validation process is *not* validating the aggregate AUI for either programs or sites, but rather is validating the frequency of error/accuracy for individual facilities—no weighting is being done, and only a sample of facilities is being used for this latter purpose.

The components of the AUI are operational net occupiable square footage (numerator) and the total net occupiable square feet not in decontamination and decommissioning (denominator) (D&D). Operating gross square footage is equal to net occupiable square footage multiplied by the building status utilization (a percentage). Special attention must be paid to the building status since it can be mis-coded because it has not been recently updated or has an inaccurate usage code (building code).

Asset Utilization Index

What to check	The AUI listed for the building
Element (tab name)	Report generated
How to check	From source documents
Why checked	If not listed or different from the value separately calculated, is there a reasonable explanation? (Explain in the notes section of the checklist.)

¹¹ 41 CFR, Chapters 101 and 102 (reference d), and Executive Order 12512, Federal Real Property Management Survey (reference s).

Net Occupiable Square Footage

What to check (first)	Net Occupiable Square Footage of the sample building
Element (tab name)	PBLD_NET_OCC_SQFT (Building dimensions tab)
How to check	From source documents
Why checked	This data element—which represents the gross square feet (less common areas such as bathrooms, stairways, elevator shafts, corridors, lobbies, equipment rooms, janitor rooms, pipe and vent shafts, exterior walls, and telephone closets)—is a primary component of the calculation of AUI. It is also known as usable space. For the purpose of FIMS validation, we want to estimate the frequency of error of this important primary data element.
What to check (second)	Actual Net Occupiable Square Footage of the sample building. This is the gross square footage of the building (total floor area, exterior wall to exterior wall, of a building in square feet) ¹² minus the common areas described above.
How to check	There are several means of checking net occupiable square footage: <ol style="list-style-type: none">1. Desk analysis—Using BOMA measurements previously provided, determine and record the gross square feet of the sample building.2. Desk analysis—Using engineering drawings, including any as-built drawings, that give the dimensions of the building.3. Physical survey—Physically measure the dimensions of the building, exterior wall to exterior wall for each floor, as well as the common areas that are excluded. Although it can be quite time-consuming, this method, if done correctly, provides the most accurate data.
Why checked	This is a primary component of the calculation of AUI. For the purpose of FIMS validation, we want to estimate the frequency of error of this important primary data element.

Gross Square Footage

What to check	Gross square footage of the sample building
Element (tab name)	PBLD_GROSS_SQFT (Building dimensions tab)
How to check	From source documents
Why checked	Net occupiable square footage, a primary component of the calculation of AUI, is the gross square footage minus the common areas. A quick check of GSF provides an opportunity for spotting potential errors (such as gross square footage listed as less than net occupiable square footage).

¹² See Note 1, p. A-19.

Building Status

What to check (first)	The building status listed for the sample building
Element (tab name)	PBLD_CMST_STATUS (Building info tab)
How to check	From source documents
Why checked	<p>This represents the programmatic intentions as well as the physical/operational status of the building. If correctly coded, it indicates whether the facility should be considered in the sample. Status should reflect one of the following four codes:</p> <ol style="list-style-type: none"> 1. Operating—a facility required for DOE's current and ongoing needs and responsibilities. 2. Operational standby—future programmatic use of the facility (other than cleanup) expected. 6. Operating pending D&D—facility transferred to the programmatic office or organization responsible for D&D activities. 7. Operating under an outgrant—facility used by another party through means of a lease, easement, license, or permit. <p>Any code other than 1, 2, 6, or 7 indicates a facility that should not be in AUI calculations.</p>
What to check (second)	Actual building status listed for the sample building
How to check	From on-site verification
Why checked	<p>Status should reflect one of the following four codes:</p> <ol style="list-style-type: none"> 1. Operating—a facility required for DOE's current and ongoing needs and responsibilities. 2. Operational standby—future programmatic use of the facility (other than cleanup) expected. 6. Operating pending D&D—facility transferred to the programmatic office or organization responsible for D&D activities. 7. Operating under an outgrant—facility used by another party through means of a lease, easement, license, or permit. <p>Any code other than 1, 2, 6, or 7 indicates a facility that should not be in AUI calculations.</p>

Status Utilization

What to check (first)	Status Utilization for specific building
Element (tab name)	PBLD_PERCENT_UTILIZATION (Building info tab)
How to check	From source documents
Why checked	A percentage, this indicates the portion of the facility's net square feet that is utilized when the building status is "Operating."
What to check (second)	Actual Status Utilization for specific building on-site
How to check	From documented site walkthroughs, activity "rent" payments, or surveys, such as CAIS, within the past 5 years.
Why checked	Validate accuracy of FIMS entry. Facility utilization often changes over time and needs to be periodically reexamined.

Building Status Code Date

What to check	Status Code Date for building
Element (tab name)	PBLD_CMST_STATUS (Building info tab)
How to check	From source documents
Why checked	Codes that have not been updated recently are more prone to error.

Excess Indicator

What to check	Excess Indicator code
Element (tab name)	PROP_EXCESS_IND (Prop info tab)
How to check	From source documents indicating (yes/no) whether the field office or site has designated the property as excess now or in the future.
Why checked	To filter out excess buildings, those not designated for use. It also indicates whether FIMS is filtering out excess in calculation of AUI. Facilities that are not utilized may be designated as excess, or on-site staff should offer an explanation.

Excess Year

What to check	If excess indicator is “yes,” date that property was declared excess
Element (tab name)	PROP_EXCESS_YR (Prop info tab)
How to check	From source documents
Why checked	To indicate how long the facility has been excess and not utilized. Dates more than 5 years old indicate an increased chance of error.

EVALUATION AND REPORTING

Upon completion of the validation process at all selected sample sites, the validation results, or the statistical frequency of error (material variation), are to be rolled up to the site level and reported for each metric. The FIMS validation and reporting template, and ultimately the FIMS web application, are able to automatically calculate the rolled-up site validation results. The site results can then be rolled up into a program-level result and, finally, a DOE-wide result, by the FIMS web application.

The template requires listing a “most likely cause of variance” for each FIMS element for a sampled facility at a site that varied from source documents by a material (10 percent) amount. These causes should then be categorized and analyzed to differentiate between individual or one-time causes of variance and potential systemic causes of variance.

Chapter 4

Pilot Observations

OVERVIEW

This chapter contains the results from the two pilot sites: Brookhaven National Laboratory and Morgantown National Energy Technology Laboratory. Brookhaven is part of the Science (SC) program, and Morgantown is part of the Fossil Energy (FE) program. Both sites volunteered to participate in the pilot.

In addition to being in different DOE programs, the two sites differed in other important ways, which increased the lessons learned from the pilot. Those differences include the following:

- ◆ Morgantown is a relatively small site with a total of 49 facilities (buildings and trailers), while Brookhaven is much larger with more than 675 facilities, including 375 buildings and over 300 permanent trailers).
- ◆ Morgantown performs all FE work related to the site's primary mission, with one small non-DOE (Navy) facility as a tenant. By contrast, Brookhaven has multiple tenants, including 60 percent SC and about 33 percent non-DOE.
- ◆ Morgantown uses FIMS as its primary facilities database, while Brookhaven uses a site-developed facilities database.
- ◆ Morgantown uses DOE's Condition Assessment Information System (CAIS) to estimate and track deferred maintenance, while Brookhaven uses a site-specific method and spreadsheet system.
- ◆ Brookhaven has a formal quality assurance program for facilities information, while Morgantown does not.

NATIONAL ENERGY TECHNOLOGY LABORATORY

Background

The Morgantown National Energy Technology Lab (NETL) is a government-owned, contractor-operated (GOCO) site. It is operated by EGG (prime contractor) with support from multiple subcontractors for various specific functions.

FIMS is the “database of record” for facilities information at Morgantown. The EGG facilities manager inputs all data into FIMS after approval of the DOE facilities manager for both Morgantown and Pittsburgh (a sister FE site also managed by EGG). The DOE facilities manager has job responsibilities in both facilities and contracting (time and duties split between two functional offices), which provides excellent visibility into all financial and accounting aspects of changes to site facilities.

Pilot

The pilot was conducted on June 7 and 8, 2004, by Robert L. Crosslin, LMI, the independent team leader, and observed by Andrew Duran, DOE OECM. The EGG facilities manager (Bill Poffenberger) and the DOE facilities manager (Rick Price) were members of the pilot validation team. Before visiting Morgantown, we identified 23 operating buildings and permanent trailers. On the basis of the sampling method developed for FIMS validation, we selected 12 buildings and trailers to sample (see Appendix A). However, upon arrival at NETL, we found that the correct number was 48 operating buildings and trailers.¹ (We encountered a similar problem at Brookhaven; these types of discrepancies need to be resolved before full validation, as we recommend in Chapter 5.)

During the visit, we reviewed all drawings, work requests, work orders, CAIS information, and other documents that track site facilities. We also had access to the people that use and maintain these records and systems day to day.

In testing the validity of the desk-audit approach to FIMS validation, we analyzed a “sample of the sample.” We performed this analysis by taking physical measurements of the buildings and using them to validate the utilization percentages.

RPV

GROSS SQUARE FOOTAGE

All building drawings are maintained real-time in AutoCAD, which is Morgantown’s official source of data for all square footage items. We spent several hours reviewing each of the sampled building’s drawings, floor by floor and area by area, in AutoCAD in combination with a calculator to obtain total gross square footage. Both the DOE and EGG facilities managers approved all changes in gross square footage. Both gross square footage and net occupiable square feet (NOSF) from AutoCAD were identical to those in historical reports and our previous validation snapshot of FIMS, from December 4, 2003. (See the discussion of NOSF later in this chapter for additional information.)

¹ Two structures are connected and operate as one facility, so the correct number could be 49.

RPV UNIT COST

Morgantown uses the RSMeans models within FIMS for all RPV calculations. It does not adjust the RPV unit cost values assigned by FIMS. This is a constant for each RPV_Model_Type (a FIMS variable), which Morgantown uses. As a result, the team only had to validate the GSF.

GEOGRAPHIC COST FACTOR

The geographic cost factor for Morgantown is a constant and provided automatically by FIMS, so this item did not need to be validated.

SITE FACTOR

The site factor is also a constant for Morgantown and provided automatically by FIMS, so it did not need to be validated.

RPV CALCULATION

Morgantown uses the RSMeans models within FIMS for all RPV calculations. It does not adjust the RPV cost values calculated by FIMS. Since FIMS is the database of record for NETL facilities, Morgantown had no material variances in GSF.

ACI

RPV

All validated as accurate with no variances; see the discussion under RPV.

DEFERRED MAINTENANCE COSTS

Morgantown and its contractor, EEG, have a formal process for conducting and documenting condition assessments, estimating costs for needed maintenance and repair, developing needed work requests for approval, and documenting deferred maintenance in FIMS. This process is well documented and produced results that were accurate and easy to validate. The EEG facilities manager inspects all facilities in detail on a 2½- to 3½-year cycle. Conditions that arise outside those inspections are also documented. EEG submits written work requests to DOE for approval. CAIS is used to estimate and document the costs for all work requests. When a work request is approved, it becomes a work order. Disapproved work requests are not entered into FIMS as deferred maintenance, even though the deficient conditions exist—they will recur as work requests the following year. Only the cost of approved work orders that are not completed by the end of the fiscal year in which the work should be finished are entered into FIMS as deferred maintenance. The cost of the work order is removed from deferred maintenance in

FIMS when the work order is completed and the EGG facilities manager has inspected and signed off on the work.

Although most DOE sites use CAIS to estimate the costs of repairs and maintenance, the current version of CAIS does not keep historical data or backups. Therefore, to validate the deferred maintenance costs in our December 2003 sample, we used the current deferred maintenance costs for a facility less all deferred work orders that had been entered since that date and added back all work orders closed or completed, also since that date. This reconciliation was further complicated by the fact that the site factor for Morgantown had changed since December 2003. (In striving to perform this validation, we concluded that CAIS needs some type of backup or archiving capability, with detail and summary reports as of month-ending dates.)

Again, we found that Morgantown had no material variances in deferred maintenance between its source documents and systems and FIMS. Since there were no variations for RPV, both FCI and ACI had no material variances for any sampled facilities at Morgantown.

Morgantown representatives raised the issue that the current contract with EGG calls for individual subcontractors to perform specific facility repairs. Unfortunately, those repairs are resulting in actual unit costs that are at least two times those generated by CAIS.

ACI CALCULATION

All validated as accurate with no material variances.

AUI

GROSS SQUARE FOOTAGE

See discussion of gross square footage validation under RPV.

NET OCCUPIABLE SQUARE FOOTAGE

Between our December 2003 snapshot and the validation site visit, Morgantown changed its interpretation of NOSF.² Previously, Morgantown had only counted space as “occupiable” if it could be used as an office or lab. It ignored, for example, square footage used by computers and servers, along with other usable or assignable space. The new (and more accurate) criteria now count these types of usable and assignable spaces. We validated NOSF under both definitions for all sampled facilities and found no variances.

² Telephone discussions with Andy Duran of OECM.

BUILDING UTILIZATION

Building utilization is the only FIMS data element that did not have source documentation, so it was impossible to conduct a “desk validation.” Morgantown’s current procedure is to annually request, usually at the end of the fiscal year, a utilization percentage from the space manager or building custodian. Sometime the request is made via e-mail, but no records are available. Morgantown plans to change this procedure to make all requests via e-mail and establish a paper file. We validated this data item during our sample-of-the-sample physical validation and found no material variances (see Sample of the Sample).

“Utilization” routinely encompasses a broad working definition. Some space managers have interpreted it to include space that is not currently being used but is planned for use, such as a project approved and pending funding or equipment. It also can include space that is being renovated.

Sample of the Sample for Physical Inspection and Measurement

In determining whether a desk validation was a valid approach, we sampled the sample and then physically validated the key data elements. After selecting every other facility on the sample sheet, or six facilities, we performed the following validation activities:

1. Requested and received AutoCAD printouts.
2. Measured each facility for gross square footage, and toured each facility to determine utilization percentages.
3. Site points of contact measured the first floor, outside footprint, while the independent team lead took the readings of the tape. We then checked the readings against the CAD drawings. We found no material variances (other than a few inches because of slack in the tape over 50 to 150 feet) between the physical measurements and the CAD drawings, which we had already validated against FIMS.
4. We toured all floors of each facility to estimate their actual utilization. The FIMS utilization percentages for most of the facilities were 100 percent, one was 80 percent, and one was 70 percent. Based upon our visual inspections, we found no material variances with FIMS.

BROOKHAVEN NATIONAL LABORATORY

Background

Brookhaven National Laboratory (BNL) is also a GOCO facility, operated by Brookhaven Science Associates. The installation was formerly Camp Upton, a World War I and II mobilization site. In the late 1940s, it was assigned to the Atomic Energy Commission and subsequently to DOE. It has the look of an Army installation, including extensive use of “World War II wood” that has been maintained and improved over the years. BNL also contains a few historic sites, such as a WW I ammunition storage building for which BNL is initiating disposal discussions with the New York State Historic Properties Office (SHPO).

BNL supports numerous activities for both DOE and other government agencies and universities. Approximately 60 percent of the site supports SC, with an additional 5 percent supporting other DOE activities. The remainder supports other agencies, such as the National Institutes of Health and National Aeronautics and Space Administration, and universities performing research activities funded by National Science Foundation (NSF) grants. These agencies are charged full cost for support (in the form of rent), including facility maintenance and repair.

Pilot

The pilot was conducted June 23-25, 2004, by Jonathan P. Adams, LMI, the independent team leader, and observed by Andrew Duran, DOE OECM. BNL has approximately 375 buildings and more than 300 permanent trailers. Our initial sample of 25 facilities included 12 trailers. We revised that sample set before our visit after a discussion with DOE and BNL representatives. The BNL representatives thought that the sample should not include trailers because BNL uses that designation liberally by “putting numbers on everything” to include many non-structural items such as conex containers. Thus, we revised our sample to consist of 25 buildings, 7 of which we physically inspected and measured. (However, the issue of trailers at DOE sites still needs to be resolved before full implementation.)

The pilot provided additional information on how BNL determines the application of RPV, AUI, and ACI for FIMS. We began the desk validation by reviewing the processes that BNL uses to determine values for RPV, gross square footage, deferred maintenance, and NOSF. BNL provided drawings and FIMS data sheets for every sample building and explained how the data were determined and entered into FIMS. BNL maintains a separate and locally developed database—Facilities Management System, or FMS—to track facilities information and automatically upload data into FIMS. It also uses a separate spreadsheet to track all deferred maintenance information from which FIMS deferred maintenance data are extracted.

BNL charges “rent” to all activities. It applies three separate space charges on the basis of type of facility: lab or office, high bay or industrial, or storage facility or warehouse. Maintenance to support experiments (for example, NSF-supported university research) is also a direct charge. This landlord-tenant relationship, with rent payments for space occupied, is used to manage the space while providing an economic discipline to facility management decisions.

Our physical validation consisted of measuring the seven randomly selected sample buildings and comparing four of them with the deferred maintenance cost data in BNL’s local spreadsheet database. Throughout this validation, we found no material variances between FIMS and source documents, systems, and processes for the sampled facilities at BNL.

RPV

BNL calculates RPV on the basis of the facilities that it currently maintains, so the type of facility currently used for an activity often does not match the code or requirements of a new structure for the same use. BNL does not use FIMS calculations to determine its RPV. It customized its calculations for each facility following a major initiative of early 2003 in which an architect-engineer (A-E) calculated the RPV for all BNL facilities. These calculations were based on RSMeans, but they also included the use of substantial professional judgment, such as recognizing that a “college lab” does not have the same cost as an “industrial lab.”

BNL tracks the usage of its facilities down to individual rooms because many facilities have multiple users and uses. The different uses may combine administrative offices with industrial production bays. The process of using individual rooms as the basic unit in RPV calculations, while time consuming, produces more accurate results because they are based on the current usage of specific space. BNL intends to repeat these calculations every few years, which should refine the process and shorten the learning curve each time. The process is being set so that current RSMeans values can be imported for updates.

However, BNL through its own quality assurance process identified several errors in its customized calculations, including double counting of general contractor and A-E fees, using the wrong RSMeans books, and failing to apply the appropriate cap to the fully burdened cost on some high-cost facilities.) While many of these errors were offsetting, BNL is working to correct them.

GROSS SQUARE FOOT

All building drawings are maintained in AutoCAD, which contains software that automatically calculates gross square footage. BNL conducts periodic quality assurance of its documents. AutoCAD automatically updates square footage figures for FMS and FIMS. BNL compares the figures monthly to ensure continued accuracy; it attributes small differences to rounding.

RPV UNIT COST

BNL used RSMeans data (not the DOE RSMeans estimating models) for its customized calculations of RPV.

GEOGRAPHIC COST FACTOR

The geographic cost factor is a constant for BNL and is provided automatically by FIMS. However, according to the process that BNL uses for valuating RPV, this factor was already worked into the customized calculation.

SITE FACTOR

The site factor also is provided automatically by FIMS. However, according to the process that BNL uses for valuating RPV, this factor was already worked into the customized calculation.

RPV CALCULATION

BNL does not use the FIMS-provided calculation for RPV; it prefers to use more precise calculations. Although we did not find any material variations between FIMS and the customized calculations, BNL did note that its quality assurance process had identified several errors by the individual that originally calculated the RPVs (see above). We address this issue and the means for identifying such errors in the next chapter.

ACI

RPV

See the discussion above.

DEFERRED MAINTENANCE COSTS

BNL maintains a separate spreadsheet database for tracking the various components of deferred maintenance. BNL's definition of deferred maintenance—based on the dominant workload performed at the facility—varies from that used at other DOE sites.

The spreadsheet database tracks a number of sources (columns on the spreadsheet) for maintenance. One column is used for “900 series” work, which is work supporting multiple buildings with deferred maintenance allocated across the buildings based on square footage. Other columns are based on work orders submitted, or whole building renovations needed for older facilities. However, the “Total” column is not simply a summation of columns (whole building renovation captures submitted work orders and represents in these cases the total deferred maintenance for one of these buildings). BNL facility engineering representatives

periodically meet with facility inspectors to review the database for double counting or overlapping entries.

The deferred maintenance spreadsheet database is a dynamic system. The values are constantly updated as new information comes in, so they justifiably vary from the “snapshot in time” used in FIMS. For the facilities selected for a physical verification of the desk validation, the BNL staff walked us through their documentation to show where the values originated, thus reconciling the FIMS snapshot values with their sources.

BNL has approximately \$200 million in deferred maintenance, of which \$130 million is for major rehabilitation of 11 large labs that are roughly 40 years old. While the buildings are structurally sound, the interiors and the systems (electrical, mechanical) need replacing.

ACI CALCULATION

Deferred maintenance is the critical variable for ACI. The data are uploaded directly from BNL FMS into FIMS, so we found no discrepancies between the source documents and FIMS.

AUI

GROSS SQUARE FOOTAGE

See the discussion above on gross square footage.

NET OCCUPIABLE SQUARE FOOTAGE

BNL manages facilities at the room level, so its database contains information on more than 14,000 rooms. Each room is assigned a code that indicates the type of space, including common space that is subtracted from gross square footage to calculate NOSF. FMS automatically subtracts common areas from gross square footage, providing NOSF for each facility.

BUILDING UTILIZATION

Of the 14,000 rooms on BNL’s database, each is charged to an activity with a monthly rent payment, or designated as either “common area” or vacant. Vacant space, which generates no rent, provides an incentive for finding uses for the space. Rather than use a subjective evaluation for occupancy or usage, BNL operates as a “landlord” and relies on rent charges to determine occupancy. BNL does not question the positions of the space managers. If a manager is willing to pay rent for a room, BNL assumes it is occupied. All billings are made monthly, so building and room usages are continuously updated.

Sample of the Sample for Personal Inspection and Measurement

We visited seven facilities from the sample to verify that the desk validation was accurate. We measured the outside dimensions of the first floor and then checked the readings on the tape against those in the CAD drawings. Other than very minor variances, which were often caused by slack in the tape, we found only one error: the length of Building 966 was recorded as 43 feet, but it was actually only 40 feet. We concluded that even this difference was not a material variation.

Chapter 5

Conclusions and Recommendations

OVERVIEW

This chapter presents our conclusions and recommendations for a full implementation of FIMS validation. The objectives of the pilot site validations were to test the planned FIMS validation procedures and determine whether the proposed approach was a valid and feasible method for assessing the accuracy of FIMS data. From the results of the pilot validations, we conclude that the approach is valid and feasible, and recommend that it be followed for the full validation at all sites. We also recommend several adjustments to the approach that should improve its use during the full-scale validation.

ONLINE SAMPLING, DATA COLLECTION, AND RESULTS

The full-scale validation process will be most efficient, as well as most timely, if much of it is conducted online. An online process would help to ensure consistency of procedures across sites and programs, support quick and automated sampling of facilities from the online database close to the date of the site validations, and provide for real-time and automated calculations of all results. We therefore recommend incorporating the following steps into an online validation process:

- ◆ Identify facilities from which to draw the validation sample.
- ◆ Select facilities at random for validation.
- ◆ Identify “outliers” to be added to the randomly sampled facilities.
- ◆ Prepopulate, as much as possible, the data collection form for each sampled facility.
- ◆ Record source validation data in the data collection form
- ◆ Calculate results by facility, site, program, and DOE.

WHEN AND HOW TO SAMPLE

One of the important lessons from the pilot validations was that the FIMS values for a facility, as well as the source documents and systems that feed FIMS, can change daily. The primary exception to this lesson is deferred maintenance, which

is typically updated in FIMS only at the end of each fiscal year. Because of the dynamics of the FIMS data, the timing of certain steps in the validation process is critical to minimizing the effort required to carry out the process. In particular, the sample needs to be drawn within a few days of the actual site validation to avoid extensive reconciliation between FIMS values for sampled facilities at the time the sample is drawn and later values during the site visit.

Algorithms (such as random number generators) and automated procedures (such as determining sample sizes and applying randomly generated numbers to building numbers for sampling facilities at a site) should be incorporated into the web-based version of FIMS. This sampling could generate a list of facilities, along with their current FIMS values, a day or two before the validation. The FIMS values for the sampled facilities could also automatically populate the data collection and input form described below.

We continue to recommend simple random sampling, without weighting by square feet or any other factor. The pilot sites confirmed our hypothesis that neither size of facility nor any other factor had any correspondence to the probability of an error in the three key FIMS metrics or their data elements. In the future, should a factor surface (for example, through analyzing sources of errors in the full-scale validation), the standard and procedures developed do not need to be changed—the only change required is relatively easy, adjusting the online sampling algorithm to weight the sample by square feet or other factor.

DATA COLLECTION AND INPUT FORM

For the two pilot sites, we used a simple spreadsheet application (Appendix D), which worked well and requires little change in format for the full-scale implementation. However, the spreadsheet was a standalone application at each site, and did not aggregate results for all facilities at a site, let alone aggregate results for an entire program or all of DOE. The full-scale validation process should include a database application (in Access or Oracle) incorporated into the web-based version of FIMS. This database application should perform three primary functions:

1. Automatically populate the form with current FIMS values for the sample facilities being validated.
2. Capture source data values, with appropriate “screening” functions, from the validation process.
3. Automatically calculate variances by facility and site.

CALCULATION OF AGGREGATE PROGRAM AND DOE RESULTS

A separate database application, as part of the web-based version of FIMS, should provide real-time data and reports by program and for all DOE. The data and reports should include the following:

- ◆ Aggregate statistics for each metric, by program and DOE-wide, noting the number and percentage of sites with completed validations.
- ◆ Reports on “likely sources of material variation” by metric, sorted by frequency of occurrence.

Composition and Training of Validation Team

Initially, we believed the staff at the individual sites, with assistance from an outside observer from another site or DOE headquarters, could perform the validations. However, experience at the pilot sites caused us to change this belief. Consistency in interpretation and implementation is critical to generating valid results across all programs and sites. In addition, trained and experienced team leads should be able to carry out the validations in a shorter period of time, resulting in lower cost. For at least the first round or year of validations, the lead team member for each validation should be well trained and experienced, and not a staff member of the site being validated. (Representatives from the pilot sites suggested this approach.) OECM should establish a training program for the team leads.

The validation team should have advance contact with, and the support of, all local quality assurance teams at each site. However, the validation team should lead the site validation. Representation on the validation teams by headquarters, and possibly the servicing centers, would ensure consistency in the validation process and interpretation of results, as well as promote consistency in the understanding and application of DOE policies.

Agenda and Procedures for Validation at a Site

From our experience at the pilot sites, the validation process should adhere to the following steps, each undertaken for all sampled facilities concurrently:

1. Introduce team members, review list of sample facilities, and identify local systems and people needed at various times during the validation process. Clarify any prevalidation questions and issues.
2. Validate gross square footage and NOSF, typically using real-time CAD programs and calculators as necessary. If the drawings are not current in the CAD application, review either drawings or physical measurements as deemed necessary by the team lead.

-
3. Validate other variables in the RPV metric.
 4. Review CAIS and work-order request documents for validating deferred maintenance costs and the resultant ACIs for each facility.
 5. Review source documents (possibly including physical observations) for AUI.
 6. Review validation results with the site staff, complete notes, and agree on follow-up action items to clarify any unresolved data items or issues.
 7. Conduct out-brief and prepare validation report to be distributed to appropriate parties and posted as an attachment to the data input and collection form at a later date.

Pending Issues

Several issues arose during validation at the two pilot sites that need to be resolved before full-scale implementation of the process:

- ◆ Clarify the definition of NOSF.
- ◆ Clarify policies and procedures for coding trailers as “permanent” or “temporary.”
- ◆ Clarify procedures for validating deferred maintenance, which is calculated at end of fiscal year, but validation occurs at other times.
- ◆ Improve consistency between installations on definitions (for example, deferred maintenance) and how data are derived (for example, disapproved work orders are counted at some sites, but not at others).
- ◆ Resolve differences between CAIS estimates and experience with local contractors in costs of repairs.
- ◆ Clarify policies on “source documents” for percent utilization.

Next Steps

Considering the significantly different approaches that BNL and Morgantown use to manage their facilities, DOE may find it worthwhile to examine and document some of the “best practices” already performed by certain sites or programs. These best practices can then be formalized into consistent policy guidance that defines the intended meaning, interpretation, and application of the key metrics within FIMS. We’ve documented some best practices in Appendix B, but it is only a beginning to what needs to be a comprehensive DOE-wide assessment.

In addition, we recommend DOE develop an automated validation application as an integrated part of the new FIMS web application.

Finally, we recommend proceeding with Phase 3 of the original task, which calls for identifying methods for OECM oversight and recommendations for built-in controls to drive better FIMS quality.

Appendix A

References

We used the following documents in the course of this preparing this report.

Department of Labor, *Employment and Training Administration, TAA Data Validation Handbook*, January 2004.

Lawrence Livermore National Laboratory, *DOE FIMS Database Quality Assurance Plan*, 2003.

Naval Facilities Engineering Command, *Facility Support Contract Quality Management Manual*, MO-327, Change 94-01, July 1, 1994.

Pantex, *FY2004 FIMS Quality Assurance Plan*, 2003.

TetraTech, Inc., *Facilities Information Management System (FIMS) Quality Assurance Plan for the Portsmouth Gaseous Diffusion Plant*, February 2003.

U.S. Army Corps of Engineers, *Chemical Quality Assurance for HTRW Projects*, EM 200-1-6, October 10, 1997.

U.S. Army Corps of Engineers, *Requirement for the Preparation of Sampling and Analysis Plans*, EM 200-1-3, February 1, 2001.

U.S. Army, *Reliability Primer for Command, Control, Communications, Computer, Intelligence, Surveillance, and Reconnaissance (C4ISR) Facilities*, TM 5-698-3, July 10, 2003.

U.S. General Accounting Office, *Applied Research and Methods, Assessing the Reliability of Computer-Processed Data*, GAO-03-273G, October 2002.

USEPA–New England, Region I, *EPA–New England Data Validation Functional Guidelines for Evaluating Environmental Analyses*, December 1996.

Appendix B

Best Practices

LMI reviewed similar data validation practices of several public and private sector organizations. We found that each had benefits, depending on the purpose and desired outcome of each validation. Although many of the practices do not directly apply to our study, they do provide information on the importance of data validation and key validation concepts. The following are key practices and concepts we've discovered through our research:

- ◆ Using commercial software to validate data
- ◆ Assigning surveillance levels to a data element on the basis of its critical or importance rating
- ◆ Assessing reliability to determine mission success, safety, availability, and life-cycle costs
- ◆ Developing project-specific sampling and analysis plans (SAPs) for the collection of data
- ◆ Preparing specific checklists for field sampling
- ◆ Providing instructions on resolving or determining sources of error, frequency and duration, and required training and certifications
- ◆ Weighing certain data elements to determine the level of effort for data validation
- ◆ Having an individual who did not perform the original input verify data entry
- ◆ Verifying data by validation sampling plan and schedule
- ◆ Validating different sections of data each year to ensure all data are validated every few years.

REVIEW SUMMARY

The following are the documents we reviewed:

- ◆ Department of Labor, *TAA Data Validation Handbook, Employment and Training Administration*, January 2004. This handbook provides instructions to states for reviewing the accuracy of their Trade Act Participant records to the Employment and Training Administration. TAA uses commercial software to conduct its data validation, and this handbook is a systematic guide on using the software for validation procedures, record layout, performance measure specifications, and sampling and error rate estimation.

The data elements the software validates are broken into three factors: feasibility, risk, and importance:

- Feasibility concerns data elements of measure and does not include data elements that are self-reported, such as race, ethnicity, and gender.
 - Risk concerns data elements involving human judgment.
 - Importance concerns data elements that are selected for validation on the basis of their importance to the integrity of the individual participant records for generating performance.
- ◆ Naval Facilities Engineering Command, *Facility Support Contract Quality Management Manual*, MO-327, Change 94-01, July 1, 1994. This study provides naval shore activities guidance on obtaining quality public works support services through facility support contracts. The examines data using three levels of surveillance methods: 100 percent inspection, planned sampling, and random sampling.

The study assigns the level of surveillance to a data element on the basis of its critical or importance rating. Data elements of the highest importance are 100 percent inspected for output accuracy, while those of only moderate importance are planned sampled, calculated by dividing the total number of observed defects by the total population. The data elements of least importance are randomly sampled.

The 100 percent inspection ensures that the data entered are complete and accurate. This method is used when the importance of the accuracy of the results is high and the population is small. The drawbacks are cost and time, if the population is large. Planned sampling works best with a small population and when only defined areas need sampling. Random sampling is the least expensive of the three and is best used when the population is homogenous and time is limited.

- ◆ U.S. Army Corps of Engineers, *Reliability Primer for Command, Control, Communications, Computer, Intelligence, Surveillance, and Reconnaissance (C4ISR) Facilities*, TM 5-698-3, July 10, 2003. This primer introduces and overviews reliability. It provides the tools necessary to evaluate a product's performance over a period. Ensuring that a system does not fail before its expected lifetime has some importance. Reliability is a basic driving factor for support requirements. This primer discusses the importance of reliability in regard to mission success, safety, availability, and life-cycle costs.
- ◆ U.S. Environmental Protection Agency (EPA)–New England, Region I, *EPA–New England Data Validation Functional Guidelines for Evaluating Environmental Analyses*, December 1996. These guidelines provide technical direction to ensure that only data of known and documented quality are used in making environmental decisions. Region I uses this for standard operating procedures, and it provides functional guidelines for three areas of environmental data validation. The data validation process, the same for all three areas, includes the following:
 - Checking data completeness of all required samples, documents, and other items
 - Evaluating all QC checks and analytical procedures, including preservation techniques and quality control sample results
 - Examining the raw data in detail to verify the accuracy of the results reported by the laboratory, checking 10 percent of the reported samples for error. Unless errors are found, no further sampling is done. The benefit of sampling 10 percent of the population is that it provides an additional low cost and time saving quality check to ensure that the validator followed each step.

The data are summarized in a six-part data validation report; the guidance describes each section.

- ◆ U.S. Army Corps of Engineers, *Requirement for the Preparation of Sampling and Analysis Plans*, EM 200-1-3, February 1, 2001. This guidance provides useful information on the preparation of project-specific sampling and analysis plans (SAPs) for the collection of environmental data. In addition, it provides default sampling and analytical protocols, which can be applied to project-specific data. The data used for this study are separated into two parts: (1) field sampling plan (FSP) and (2) quality assurance project plan (QAPP). The FSP addresses field activities, including all aspects of sampling, drilling, monitoring well installation, and any field data gathering activities. The QAPP addresses the data quality objectives, analytical methods, specific quality assurance (QA) and QC activities,

laboratory requirements, and data assessment activities designed to achieve quality data goals.

The guidance provides detailed instructions on the planning, format, and content of SAPs and defines the roles and responsibilities of those working on the SAP. It is helpful to both data users and implementers who are looking to ensure that quality data are generated.

The data verification and validation of this study refers to US Army Corps of Engineers, EM 200-1-6, *Chemical Quality Assurance for HTRW Projects*, October 10, 1997. This manual provides specific guidance, procedures, criteria, and tools for chemical implementation of Corps of Engineers hazardous, toxic, and radioactive waste projects. It provides very specific checklists for field sampling and includes instructions on resolving or determining sources of error, frequency and duration, and required training and certifications.

- ◆ U.S. General Accounting Office, Applied Research and Methods, *Assessing the Reliability of Computer-Processed Data*, GAO-03-273G, October 2002. This guidance explains the meaning of data reliability and provides a framework for assessing the reliability of computer-processed data. It gives some insight in determining whether the data available are useful and the level of risk.

COMMON ELEMENTS

The reports and our analysis both stress the level of importance in acquiring complete and accurate data in order to conduct data validation that will produce useable information. The reports that we reviewed all agree that the first step in validating data is that all data need to be reviewed for completeness, thus requiring the data validation to contain all necessary information presented in the format requested.

Many of the reports are guidance documents or handbooks, and each goes into specific detail on the appropriate methods for reporting, who is responsible for the reporting, and the corrective measures when the reporting fails. The key item is that accurate and detailed data reporting is the key to successful data validation.

The *TAA Data Validation Handbook* best resembles our approach to data validation. It looks to determine the accuracy of the data by evaluating three specific areas. In this report, we are looking to develop our standard to include three key FIMS metrics (RPV, ACI, and AUI), all entered by the field and all having significant influence on the financial obligation of the results.

The *Facility Support Contract Quality Management Manual* approach—weighting certain data elements that will determine the level of effort for data

validation—is also pertinent. DOE currently maintains 22 critical data elements that have higher importance as data needed or required to be entered by the sites.

ALTERNATIVE APPROACHES

The *TAA Data Validation Handbook* looks beyond the capabilities of conventional data validation, implementing a software system that conducts the validation for them. Microsoft Access supports the validation software, developed by Mathematic Policy Research, Inc. The software helps the Employment and Training Administration meet its performance goals by comparing the values calculated by states to the values of selected data elements. The selected elements confirm the accuracy of the reported data through a comparison conducted by the staff.

Assessing the Reliability of Computer-Processed Data provides instructions on conditions that require a data reliability assessment. An alternative approach would be to consider conducting a data reliability assessment to verify that the data in FIMS can support any finding that will result in conclusions or recommendations.

CURRENT QA PROCESS

We reviewed the QA plans for three DOE sites and found that each contains similar purposes and goals, but that each has minor difference directly related to the type of facilities maintained and the process used to validate site information. Although each site validates the data annually, significant differences may affect the site's quality of data:

- ◆ *Portsmouth Gaseous Diffusion Plant (PORTS)* uses an onsite Windows-based database to track real property information and transfers the information to the FIMS database for HQ reporting. Bechtel Jacobs Company (BJC) is the supporting contractor for maintaining and reporting of data and verifies data accuracy using physical verification and management overview of the activities involved with developing facility and maintenance data or directly inputting the data. PORTS has BJC subcontract the data entry support, and the responsibilities of this contract include data entry, retrieval of data for report generation, accurate data entry, documentation of changes, maintenance, and storage of FIMS reports. Data entry is 100 percent verified by an individual who did not perform the original input.
- ◆ The *Lawrence Livermore National Laboratory (LLNL) QA Plan* documents the aspects of quality assurance including organization, responsibilities, training, validation, and certification that LLNL information is accurately collected and reported. Data verification is ongoing at LLNL and is done in accordance to their validation sampling plan and schedule.

A minimum of 10 percent of all facilities is randomly surveyed annually as part of a formal survey process to verify accuracy. The current process surveys different facilities each year to insure that all facilities are visited a minimum of once every 10 years.

- ◆ The *Pantex Plant QA Plan*, managed by their M&O contractor, BWXT Pantex, LLC (BWXT Pantex), is similar to both PORTS and the LLNL plan. The overall responsibility for managing and maintaining FIMS falls under the Readiness and Technical Base and Facilities Program in the Master Site Planning Department. All data obtained from original sources are presumed valid; if the data source is received via an external source, it is reviewed for accuracy. A majority of the validation occurs during the scheduled quarterly and annual maintenance and is coordinated with its other work management systems (PassPort & CAIS). These updates provide an additional check for data accuracy. Dyncorp conducts another validation used at Pantex. Pantex also has their subcontractor, Dyncorp, conduct a 100 percent population of the database every quarter for added assurance in quality data.

Appendix C

Calculated Sample Sizes

In Table C-1, n' is the required sample size used for a site with a given number of facilities, N .

Table C-1. Required Sample Sizes

N Population of program i at installation k	n' Required sample size of program i at installation k
1	1
2	2
3	3
4	4
5–6	5
7	6
8–9	7
10–11	8
12–14	9
15–16	10
17–20	11
21–23	12
24–27	13
28–33	14
34–39	15
40–46	16
47–56	17
57–69	18
70–87	19
88–113	20
114–154	21
155–232	22
233–242	23
>242	25

Appendix D

FIMS Validation Handbook

INTRODUCTION

The purpose of this handbook is to serve as a “How To” guide for conducting a FIMS validation at a particular site. The validation process consists of three phases: initialization, validation, and evaluation and reporting. In the remainder of the handbook, we provide step-by-step instructions for performing all three phases.

OBJECTIVE

The objective of the validation is to verify the accuracy of three key FIMS metrics. For the aggregate results to be properly interpreted and useful for decision-making, it is important that the steps taken to validate each metric be followed consistently among all sites and programs. The three key metrics to be validated are:

- ◆ Replacement plant value (RPV)
- ◆ Asset condition index (ACI)
- ◆ Asset utilization index (AUI).

In validating these metrics, the following sequence of phases and steps should be used to check the FIMS data elements comprising these metrics against their source documents and to report the results.

Phase 1: Initialization

STEP 1: DESIGNATE VALIDATION TEAM

The ideal validation team consists of three people. Two should be intimately familiar with the facilities being validated, including their history and source documents. The third member should be a neutral party, who is not associated with the program or site being validated. This team member can be a representative from another DOE program, the site’s service center, or headquarters. This member serves as the leader of the validation team, and also documents and reports on the results. The responsibilities of the other two members of the team include providing access to source documents and systems, arranging meetings with site staff who create or maintain source documents, and supporting logistics and other

needs related to the validation. The use of a neutral party continues the FIMS learning process and promotes the surfacing of validation issues common to multiple sites.

STEP 2: DETERMINE VALIDATION SCHEDULE

Several weeks before validating an individual site, the team leader sends a copy of the blank FIMS Validation and Reporting Template, a proposed agenda, and a tentative schedule to the site. The leader also notifies the site that access will be needed to drawings (such as AutoCAD), work requests, work orders, FIMS, other databases used to track site facilities, Condition Assessment Information System (CAIS) reports, and staff members who routinely use and maintain these records or systems. The sample agenda (see Step 6), which has been used at the two pilot sites, is very effective and will save time and effort if it is followed.

The scheduling of individual sites can be a matter of convenience and efficiency for a validation team. The timing of visiting individual sites does not affect the aggregate accuracy of either the program or DOE. However, a program's (and DOE's) accuracy cannot be validated until all sites in the program have been visited.

STEP 3: ESTABLISH FIMS DATA BASELINES

Before each validation, DOE headquarters (or a FIMS contractor) pre-populates the FIMS Validation and Reporting Template with data using a random sample of facilities to be validated. These data include the facility name and identification number, along with a date stamp; they also serve as the baseline data to be validated. To ensure integrity in the validation process, the site validation team must use the data recorded in FIMS as of this date, and no data from FIMS or any other facilities databases after this date. Because FIMS is a dynamic system—values for individual facilities can change daily—the baseline data should be downloaded within 3 to 5 days of the actual site validation date. This practice avoids time-consuming reconciliation between the baseline and current data.

STEP 4: GENERATE SAMPLE SET

- ◆ The sampling process uses an on-line application attached to the Web version of FIMS. The application first determines the sample size based on the number of facilities at the site, then it automatically generates many random numbers and concludes by applying those random numbers to the facilities at the site, which yields a sample of randomly selected facilities for validation. These steps are described in more detail below:
- ◆ Identify all facilities at the site by labeling each building with an identification number, such as MARS #001, MARS #002, and MARS #003.

- ◆ Apply filters to identify the subset of facilities to be sampled, such as buildings and real property trailers (MARS #001), no leased facilities, no personal property trailers (MARS #025), and no facilities identified for demolition or excess.
- ◆ Access the sample size lookup table to determine the required sample size for the site (see Table D-1).

Table D-1- Required Sample Sizes

N Population of program i at installation k	n' Required sample size of program i at installation k
1	1
2	2
3	3
4	4
5–6	5
7	6
8–9	7
10–11	8
12–14	9
15–16	10
17–20	11
21–23	12
24–27	13
28–33	14
34–39	15
40–46	16
47–56	17
57–69	18
70–87	19
88–113	20
114–154	21
155–232	22
233–242	23
>242	25

-
- ◆ Use a computerized random number generator (freely available) to obtain a list of random numbers equal to the required sample size (a table of random numbers from any elementary statistics book can also be used).
 - ◆ Associate the random numbers (based on their sequencing pattern) with specific facilities at the site.
 - ◆ Identify “outliers” (from parameters set by OECM) to be added to the random sample, such as facilities with exceptionally high RPV and low gross square footage or exceptionally high gross square footage and low RPV.
 - ◆ Download the FIMS values for the metrics and associated data elements for the selected facilities (plus any outliers) to the FIMS Validation and Reporting Template for the site. (The reporting templates are shown at the end of this appendix.)

STEP 5: IDENTIFY SOURCE DOCUMENTS

The site validation team uses the FIMS Validation and Reporting Template as its basic tool for the validation process. All FIMS values should be supported by source documents of such quality that they could support an audit. The local site members of the validation team should identify, locate, and put the documents in a conference room prior to the arrival of the validation team lead. These source documents could include

- ◆ construction drawings and blueprints,
- ◆ as-built drawings,
- ◆ condition assessment reports,
- ◆ work order requests and orders,
- ◆ other DOE databases (CAIS or Maximo only for FIMS validation),
- ◆ DOE-specific models and benchmarks (such as RSMeans or BOMA),
- ◆ physical measurements, and
- ◆ internal memorandums.

Many sites also maintain other (site- or program-specific) facility databases, which may contain some or all of the FIMS data being validated. These databases, like FIMS, are repositories of information, so they may not be the true source documents for the purposes of FIMS validation. As a consequence, they are not acceptable substitutes for the source documents.

Phase 2: Validation

STEP 6: KICKOFF MEETING WITH THE SITE TO FINALIZE AGENDA

Table D-2 shows a sample validation kickoff agenda.

Table D-2. Sample Agenda for Site Validation Visit

Agenda item	Allotted time	Description of validation activity
Introductory meeting	1 hour	Introduce team members, review list of sample facilities, identify persons and systems required for access at planned times, identify initial issues
Validate gross sq ft and net occupied sq ft	2-3 hours	Validate these metrics concurrently, typically by comparing FIMS values with real-time CAD systems or reports, use calculators as necessary; could involve hard-copy drawings; collect documentation
Validate other fields of RPV	2-3 hours	Validate other fields; information may come from RS Means or other sources, some may be constant across all facilities at a site; collect documentation
Validate deferred maintenance and ACI	2-4 hours	Review CAIS, recently completed construction contracts, work order request systems, and related back-up documentation; explain deferred maintenance for each sampled facility; collect documentation
Validate AUI	1-2 hours	Review source documents (including e-mails, building manager memos, and rent bills) for percent utilization; may require direct observation of facility usage; collect documentation especially local policies and procedures
Conclude and review meeting	1-2 hours	Review results with validation team and site managers, complete notes, identify follow-up site or headquarters action items to clarify, document outstanding items or issues
Prepare report	8 hours	Prepare validation report with a copy to program and site

STEP 7: UNDERSTAND AND REVIEW THE FIMS VALIDATION AND REPORTING TEMPLATE

The FIMS Validation and Reporting Template automatically calculates the metric values based on inputs from source documents, as well as the variances from the FIMS baseline metrics that are downloaded to the template for the sampled facilities. The template also adds the frequency of material variances (the number of facilities with a variance greater than 10 percent) and then calculates a site variance frequency percentage for each metric.

During the introductory meeting, the validation team reviews the template to ensure a common understanding of the required information. For each required data element, the validation team leader enters the following information into the FIMS Validation and Reporting Template:

- ◆ *Source*. The most recent source document for the data element in question.
- ◆ *Year of source*. The year the source document was created, modified, or last updated. If the source document is older than 1 year for utilization, 3 years for DM or CAS inspection if mission essential, 5 years if not mission essential, or 1 year for RPV, 1 year for site factor (from the date of the validation), then the value of that FIMS metric is “automatically in error (a material variance).”
- ◆ *Value from source*. The value from the source document that corresponds to the data element in question. In the case of square footage, it may be necessary to extract individual measurements from drawings and calculate the square footage.
- ◆ *Most likely cause of variance*. The validation reporting form calculates the difference between the FIMS and source values to determine whether it is greater than 10 percent. If it is, the validation team tries to identify the likely cause of the variance. Possible causes include
 - out-of-date (see year of source above), inaccurate, or incomplete sources;
 - sources that probably exist, but cannot be located;
 - no sources;
 - multiple or conflicting sources;
 - incorrectly calculated values; and
 - incorrectly entered values.
- ◆ *Additional comments*. The validation team uses this space to document other observations or issues regarding the validation of FIMS data for a particular facility or to further explain data sources and variances.

STEP 8: IDENTIFY AND VALIDATE REQUIRED DATA ELEMENTS FROM SOURCE DOCUMENTS

The metrics and associated data (primary and secondary) elements listed in Tables D-3 to D-5 in this section are included in the validation. Since some data elements feed into more than one metric, the validation reporting form is consistent with the sequence of data elements in the tables, which is the sequence to be followed when performing the validation. The results of reviewing the source documents for these elements are then entered into the FIMS Validation and Reporting Template.

Table D-3. RPV Data Elements

Textual name	FIMS element name
Building RPV	PBLD_BUILDING_RPV
Gross square feet	PBLD_GROSS_SQFT
Site factor	PBLD_LAB_USAGE_PERCENT
Geographic cost factor	SITE_GEOCOST_FACTOR
RPV model type	RPVM_MODEL
RPV unit cost	RPVM_UNIT_COST
Adjustment cost	CAPI_IMPROV_COST
Adjustment date	CAPI_IMPROV_DATE
Adjustment description	CAPI_IMPROV_DESC

Table D-4. ACI Data Elements

Building ACI	FIMS report
Building ACI	FIMS report
Building FCI	FIMS report
Deferred maintenance	DEFM_DEF_MAINT
Building RPV ^a	PBLD_BUILDING_RPV

^aThe results obtained in the RPV metric validation are used as input to the ACI metric validation.

Table D-5. AUI Data Elements

Building AUI	FIMS report
Asset utilization index	FIMS report
Net occupiable square footage	PBLD_NET_OCC_SQFT
Gross square feet	PBLD_GROSS_SQFT
Building status	PBLD_CMST_STATUS
Status utilization	PBLD_PERCENT_UTILIZATION
Status code date	PBLD_CMST_STATUS
Excess indicator	PROP_EXCESS_IND
Excess year	PROP_EXCESS_YR

STEP 9: DOCUMENT ALL FINDINGS

Before the team leader leaves the site, the entire validation team reviews the FIMS Validation and Reporting Template for completeness for all sampled facilities. The team verifies that all input fields contain entries and that comments exist for all data elements and metrics with material variances (errors greater than 10 percent). Copies of reports, source documents, and other notes are clearly labeled and placed in a folder or binder to remain at the site for potential reference at a later date; however, the team also receives a full set of all reports, source documents, and other notes. All outstanding issues and action items, with dates for completion, should be assigned to members of the validation team.

Phase 3: Evaluation and Reporting (Step 10)

After completing the validation process at a particular site, the validation results, or the statistical frequency of error (material variation), are automatically aggregated to the site level and reported for each metric. This aggregation is performed by the Web-based database application. If some sites perform validations before the Web application is completed, the site-aggregation procedure for each metric is as follows:

1. Count the number of sample facilities with a material variance (a FIMS baseline metric value that differs from the calculated value by more than 10 percent, using the source document calculation as the denominator).
2. Divide the number of sampled facilities with a material variance for the metric by the total number of sampled facilities. For example, if 5 facilities contained a material variance for ACI and 25 facilities were sampled, then divide 5 by 25, which yields 20 percent.
3. The following statement (inference) is then made: “The frequency of material variance (error) in the RPV/ACI/AUI metric is estimated to be X percent (with 90 percent confidence in that estimate).” In the example of 5 of 25 facilities with a material variance, $X = 20$ percent.

The template requires listing a “most likely cause of variance” for each FIMS element for a sampled facility that varied from the source documents by a material (10 percent) amount. These causes are then categorized and analyzed to differentiate among individual or one-time causes of variance and potential systemic causes of variance, particularly at the program and departmental levels. The validation team notes any significant findings about the causes of variance at the site.

The FIMS Web application version of the template automatically calculates the total program and DOE results following the same methodology described above.

The following pages contain prototypes of Validation and Reporting Templates that can be used during the validations. The gray fields are to be completed by the

validation team. These templates are sufficient for initial rollout, but we recommend that a dedicated FIMS validation application be developed as part of the new Web version of FIMS.

The categories are defined as the following:

- ◆ *Source*: Enter the type of documentation used to identify the RPV or GSF, such as site drawings, recent construction plans, human measurement, or items reported in CAIS.
- ◆ *Year of source*: Enter the date of the source document.
- ◆ *Value from source*: Enter the value of the source document. For example, the CAIS value for GSF.
- ◆ *Most likely source of variance*: Enter a possible reason for the discrepancy from FIMS.

DOE FIMS Validation
 Sampled Facilities Reporting Form
 DOE Site: Name of Site and Program

SAMPLED FACILITY 1 - RPV

FIMS Facility Identification #	pre-populated (locked cell)	
--------------------------------	--------------------------------	--

DATA ELEMENT CHECKS

Metric: RPV	FIMS Value	Source	Year of Source	Value from Source	Variance (FIMS - Source)	Material Variance (Y/N)	Most Likely Source of Variance
Building RPV	pre-populated baseline values (locked cell)	text input	number input	number input	auto-calc (locked cell)	auto-calc and fill (locked cell)	text input
Gross Square Feet							
RPV Model Type							
RPV Unit Cost							
Geographic Cost Factor							
Site Factor							

BASELINE VS. CALCULATED

RPV FIMS Value =	pre-populated baseline value	pre-populated Gross Square Feet	pre-populated RPV Unit Cost	pre-populated Geo. Cost Factor	pre-populated Site Factor
RPV Value calculated from Source Elements =	calculated value	auto-filled source value: Gross Square Feet	auto-filled source value: RPV Unit Cost	auto-filled source value: Geo. Cost Factor	auto-filled source value: Site Factor
Material Variance?	Y or N auto- calculated and auto- filled				
Most Likely sources of variance if Yes	text input				

ADDITIONAL COMMENTS

--	--

SAMPLED FACILITY 2
 (repeat for all sampled facilities at site)

DOE FIMS Validation
 Sampled Facilities Reporting Form
 DOE Site: Name of Site and Program

SAMPLED FACILITY 1 - ACI

FIMS Facility Identification #	pre-populated (locked cell)	
---------------------------------------	--------------------------------	--

DATA ELEMENT CHECKS

Metric: ACI	FIMS Value	Source	Year of Source	Value from Source	Variance (FIMS - Source)	Material Variance (Y/N)	Most Likely Source of Variance
Building ACI	pre-populated baseline values (locked cell)	text input	number input	number input	auto-calc (locked cell)	auto-calc and fill (locked cell)	text input
Building FCI							
Deferred Maintenance							
Building RPV							

BASELINE VS. CALCULATED

ACI FIMS Value =	pre-populated baseline value	pre-populated FCI	pre-populated Deferred Maintenance	pre-populated RPV
ACI Value calculated from Source Elements =	calculated value	auto-filled source value: FCI	auto-filled source value: Deferred Maintenance	auto-filled source value: RPV
Material Variance?	Y or N auto- calculated and auto- filled			
Most Likely sources of variance if Yes	text input			

ADDITIONAL COMMENTS

	Note: if RPV is found inaccurate, the ACI metric will automatically be flagged as inaccurate. Each ACI element must still be validated to determine if any additional causes of variance exist.
--	---

SAMPLED FACILITY 2
 (repeat for all sampled facilities at site)

DOE FIMS Validation
 Sampled Facilities Reporting Form
 DOE Site: Name of Site and Program

SAMPLED FACILITY 1 - AUI

FIMS Facility Identification #	pre-populated (locked cell)	
---------------------------------------	--------------------------------	--

DATA ELEMENT CHECKS

Metric: AUI	FIMS Value	Source	Year of Source	Value from Source	Variance (FIMS - Source)	Material Variance (Y/N)	Most Likely Source of Variance
Net Occupiable Square Feet	pre-populated baseline values (locked cell)	text input	number input	number input	auto-calc (locked cell)	auto-calc and fill (locked cell)	text input
Building Status							
Status Utilization							
Status Code Date							
Excess Indicator							
Excess Year							

SINCE AUI IS CALCULATED AT THE SITE LEVEL, INDIVIDUAL FACILITY AUI CALCULATIONS DO NOT APPLY

ADDITIONAL COMMENTS

	Note: if RPV is found inaccurate, the ACI metric will automatically be flagged as inaccurate. Each ACI element must still be validated to determine if any additional causes of variance exist.
--	---

SAMPLED FACILITY 2
 (repeat for all sampled facilities at site)

DOE FIMS Validation
Sampled Facilities Reporting Form
DOE Site Name: Name of Site and Program

Evaluation and Reporting

Sample No.	FIMS Facility ID #	Material Variance? (Y/N)
1		
2		
3		
4		
5		
6		
7		
8		
9		
10		
11		
12		
13		
14		
15		
16		
17		
18		
19		
20		
21		
22		
23		
24		

Total # of buildings with material variance (sum of all Y's)	
Total # of buildings sampled	
The frequency of material variance (error) in the RPV/ACI/AUI Metric at 90% confidence estimated to be (divide total number of buildings with variance by total number of buildings sampled.)	

